Universitat de Girona

NEW METHODS FOR TRIANGULATION-BASED

SHAPE ACQUISITION USING LASER SCANNERS

Josep FOREST COLLADO

Universitat de Girona

Departament d'Electrònica, Informàtica i Automàtica

Tesi Doctoral

New methods for triangulation-based shape acquisition using laser scanners

Tesi presentada per:
Josep Forest i Collado

per obtenir el títol de:
Doctor en Enginyeria Informàtica

Directors:
Joaquim Salvi i Enric Cabruja

Girona, maig 2004

Contents

List of Figures

List of Tables

Agraïments

Semblava que no podria ser, però ja hi som. Deu fer uns cinc anys, fet i fotut, que vaig començar-ho de manera seriosa, fent els cursos de doctorat del departament d'EIA. Però com tots els meus companys de feina ja deveu saber, en fa uns quants més que em bellugo per aquí.

No podia ser diferent, de qui primer me n'he recordat quan començo a escriure aquestes ratlles, és d'en Joan Batlle, qui, per allà al 93, gairebé em va fer com de segon pare. Amb en Joan hem passat bones estones: una beca *de matins*, que va resultar ser *de matí a matí...*, o quan feiem en ROGER per anar *tranquil·lament* fins a Finlàndia a una competició... d'aquell temps, encara trobo gent a Europa que sap d'on vinc, quan dic que sóc de la Universitat de Girona, o quan varem anar a Noruega, amb aquell cotxe del dimoni que haviem de pujar als ferris pitjant-lo..., en fí, si em deixo res, ja m'ho diràs. Només volia agrair-te molt totes aquelles vivències i consells, sempre acompanyades de molt bon humor. D'aquells començaments, seguint aquest ordre cronològic, també val la pena recordar a en Xevi Cufí, el meu actual company de despatx, en Joan Martí, en Lluís Pacheco, i en Miki Villanueva. A en Xevi i en Joan, us haig d'agrair molt, potser més del que us imagineu. Jo acabava de *sortir de l'ou*, com aquell qui diu, i vosaltres sempre us heu portat com autèntics companys i amics de debò, animant-me a participar en tot, encara que només fés que dir una bestiesa darrera de cada ximpleria. A en Lluís i en Miki, crec que cal recordar moltes vivències i bones estones, però com que l'espai és limitat, voldria recordar el que va ser aquell famós *estiu'93*, amb la seva culminació dinant tots junts a La Font. En Lluís encara deu buscar l'entremès, però si hi ha quelcom que se t'ha de reconèixer, és els... el, diguem-ne, coratge que vares tenir per fotre cap a 100 quilos de ferro a dins del teu R-5 i anar-te'n a Hèlsinki voltant per Rovaniemi.

Poc després vàreu anar arribant en Rafa Garcia i en Quim Salvi, i més tard, en Pere Ridao i en Jordi Freixenet. D'aquella època recordo especialment la *lluita* per una vida digna dins de la Universitat, amb aquelles reunions o *mini-assemblees* que fèiem, i aquelles trobades amb l'aleshores vicerrector de personal, i aquell famós: *sí, però si no es compleixen els acords, fotrem una vaga !!!* d'en Pere, en una trobada amb el vicerrector. En Quim, que és un dels co-directors d'aquesta tesi, és qui em va *engrescar*, i de quina manera, en aquest món del 3D. Quan feiem coses submarines, anàvem a *banyar* a Banyoles per constatar que NO ES VEIA ABSOLUTAMENT RES !!!. Per sort, finalment varem decidir fer les coses *en sec*, i la de coses que han sortit. Per fortuna ja hi hem afegit un granet de sorra, i ara, potser tornem a l'*Underwater*, no ?!? Gràcies per *engrescar-me*, encoratjar-me, recolzar-me i

sobre tot, creure en el que he estat fent, malgrat la meva desbordada bogeria.
Gràcies per ser amic meu. La cosa ha crescut, i ara ja en som uns quants
que hi estem treballant, tant des de la realització de tesis doctorals, com de
projectistes final de carrera. Per tant, vull tenir també un record pels meus
companys de doctorat, els més propers, com en Jordi Pagès, en Radu Orghi-
dan, en Carles Matabosch, en Xavi Pinsach i en Marc Bigas, tots vosaltres
fent tesis en 3D, però també aquells qui varen començar els cursos amb mí:
en Xevi Muñoz i en Xavi Lladó, que ja heu acabat i heu fet la tesi en visió,
o en Xavi Armangué, que més que una tesi en estèreo, en vares fer dues o
tres, *grapades* en el mateix llibre. Per en Xpinsach, de totes maneres, voldria
agrair-li molt especialment la seva decisió final. Va costar, però al final vas
dir que sí, i has apostat pel projecte. Estàs completament sonat !!! però jo,
i no crec que sigui l'únic, t'ho agraeixo molt.
L'altre que falta també és de la colla dels kpsdeporrks!! (ja m'enteneu), en
Carles Pous (o Wells pels amics). Ja es veu que cal estar molt sonat per
treballar en aquest departament, i en particular, en coses de làsers, càmeres,
sensors, electrònica i CBRs, alhora!!!, en fi, tiu, si no hi fóssis no seria el
mateix. Polifacètic, políglota i politècnic (de vegades fins i tot polimòrfic ...
:-))), gràcies pels cafès, tertúlies, reunions, plans, etc... gràcies per ser el
meu amic.
Arribat en aquest punt, no em puc estalviar de fer un esment molt especial.
Hi ha vegades a la vida en què els planetes del sistema solar queden alin-
eats (o desalineats) d'una forma especial. Són moments de la vida en què et
sembla que has nascut amb allò que se'n diu *una flor al cul*, per què, en el
meu cas, aquesta situació tan excepcional, es tradueix en la coneixença d'uns
estudiants, que potser no són els més brillants, però que crec que costarà
de tornar-ne a trobar d'iguals. Són gent que no només han cregut que valia
la pena ajudar-me i que han cregut en una idea que estem desenvolupant,
sinó que a més a més estan treballant de franc !! Per això vull recordar molt
especialment a l'Isidro Rodríguez, en *Sidro*, en Josep M. Teixidor, en *Tesh*,
en Jordi Madrenas, l'*Adri*, en Carles Royan, en *Xarli*, i en Frederic Garcia,
en *Fredi*. Sincerament, no sé que carai hauria aconseguit sense vosaltres, us
dec un sopar, i més.
No pas per què sigui l'últim, sino per què és una persona especial, ara li toca
a l'Enric Cabruja. La veritat és que n'hauria de dir tantes, de coses, que no
sé pas com començar. Ja deu fer uns cinc o sis anys que ens coneixem, quan
anàvem i veniem de l'Autònoma en aquells autocars mig podrits. Em vas
acollir al CNM, on vaig aprendre més coses sobre microelectrònica que en
deu assignatures de la carrera. Em vas co-dirigir el projecte final de carrera,
amb aquell sistema 3D per anar a sota aigua !!! El 3D et va *enganxar* i ara
hi estem tots embrancats. La teva manera de veure la vida, com saps, és

molt igual a la meva, i de veritat que estic molt content que siguis l'altre co-director d'aquesta tesi, però del que estic més content és de tenir-te com a amic. Gràcies, tiu.

Tampoc no me'n descuido de vosaltres, Josep Tomàs, Toni Verdú, Lluís Magí, Íngrid Azorín i Marc Rodríguez. Amb en Josep i en Toni fa molts anys que ens coneixem, molts. Encara somio amb aquells interruptors de SIMON i amb les conilleres, com deia en Raimon, tu ja m'entens. Amb en Lluís hem fet i encara fem moltes *hardwarades*, espero seguir així, per què no és fàcil trobar algú que treballi la *materia* com tu, i que a més hagi *patit* el meu primer any de *profe*. Per tot això, i molt més, gràcies a tots vosaltres. Us haig d'agraïr també al personal de l'IIIA i del departament d'EIA, la Isela Ibrahimovic, que ja no hi és, però que també ha patit la meva activitat de recerca, la Montse Vila, que l'està patint ara, i molt bé, per cert, i a en Xevi Manyer, pels seus consells i suport, i fins i tot allò que els britànics en diuen *commitment* en la seva feina. A la Sònia Soler i la Mar Ortega, que ara la feu en altres indrets d'aquesta casa, i a la Marta Geli i l'Anna Renart, que em patiu actualment, juntament amb la resta de components del departament. Els agraïments arriben aquí, no sé pas si me'n descuido d'algú. En tal cas, incloeu-vos a la llista d'agraïts, honestament. No voldria pas, però, acabar aquestes ratlles sense, no pas agrair, sino fer un reconeixement públic i tan honorífic com es mereix, a la persona més especial de la meva vida. La Maite, la meva companya sentimental, la meva parella afectiva, la persona a qui estimo i amb qui he tingut dues filles precioses (quan dormen... no... és broma), i que ens fan viure sensacions irrepetibles en tots els sentits (proveu-ho, que val la pena), sensacions que et posen a prova, especialment quan fas una tesi. Aquesta, de tesi, és tan meva com seva. És evident que tots dos hem fet el mateix nivell d'esforç. Potser en direccions diferents, però totalment complementaries. Al final ho hem aconseguit junts tots dos, i això no té preu. Sonarà a poc, però gràcies, Maite, a tú és a qui dedico aquesta tesi. Per acabar, vull tenir unes paraules pels meus pares. Ells m'ho han donat tot, i m'han ajudat i encoratjat en tot moment i en totes les decisions que he pres a la vida, i tot i que aquestes ratlles són poca cosa, com a mínim això s'ho mereixen. Encara me'n recordo quan tu, pare, em deies: *vols dir fer enginyeria ?, mira que m'han dit que és molt difícil !!*. Bé, doncs, vet aquí que ja em teniu acabant la tesi. Us agraeixo molt tot el que heu fet (i desfet) per mi. Gràcies, pare i mare.

Resum

Tradicionalment, la reproducció del món real sen's ha mostrat a través d'imatges planes. Aquestes imatges se solien materialitzar mitjançant pintures sobre tela o bé amb dibuixos. Avui, per sort, encara podem veure pintures fetes a ma, tot i que la majoria d'imatges s'adquireixen mitjançant càmeres, i es mostren directament a una audiència, com en el cinema, la televisió o exposicions de fotografíes, o bé són processades per un sistema computeritzat per tal d'obtenir un resultat en particular. Aquests processaments s'apliquen en camps com en el control de qualitat industrial o bé en la recerca més puntera en intel·ligència artificial. Aplicant algorismes de processament de nivell mitjà es poden obtenir imatges 3D a partir d'imatges 2D, utilitzant tècniques ben conegudes anomenades *Shape From X*, on X és el mètode per obtenir la tercera dimensió, i varia en funció de la tècnica que s'utilitza a tal finalitat. Tot i que l'evolució cap a la càmera 3D va començar en els 90, cal que les tècniques per obtenir les formes tridimensionals siguin més i més acurades. Les aplicacions dels escàners 3D han augmentat considerablement en els darrers anys, especialment en camps com el lleure, diagnosi/cirurgia assistida, robòtica, etc.

Una de les tècniques més utilitzades per obtenir informació 3D d'una escena, és la triangulació, i més concretament, la utilització d'escàners làser tridimensionals. Des de la seva aparició formal en publicacions científiques al 1971 [SS71], hi ha hagut contribucions per solucionar problemes inherents com ara la disminució d'oclusions, millora de la precisió, velocitat d'adquisició, descripció de la forma, etc. Tots i cadascun dels mètodes per obtenir punts 3D d'una escena té associat un procés de calibració, i aquest procés juga un paper decisiu en el rendiment d'un dispositiu d'adquisició tridimensional.

La finalitat d'aquesta tesi es la d'abordar el problema de l'adquisició de forma 3D, des d'un punt de vista total, reportant un estat de l'art sobre escàners làser basats en triangulació, provant el funcionament i rendiment de diferents sistemes, i fent aportacions per millorar la precisió en la detecció del feix làser, especialment en condicions adverses, i solucionant el problema de la calibració a partir de mètodes geomètrics projectius.

Abstract

Traditionally, the reproduction of the real world has been shown to us by means of flat images. These images used to be materialised by means of paint on canvas, drawings or the like. Today, we still see hand made pictures, by fortune, although most of the images are acquired by cameras and they are either directly shown to an audience, like in the cinema, television or photographs, or they are processed by a computer system in order to obtain a particular result, like in industrial quality assurance or bleeding edge artificial intelligence research. Applying mid-level processing algorithms, 3D images can be obtained from 2D ones, using well known techniques called *Shape From X*, where X is the method for obtaining the 3rd dimension. While the evolution to the 3D camera begun in the 90s, the techniques for obtaining the most accurate 3D shape need to be continuously improving. The application of 3D scanners has spread significantly in the recent years, specially in fields like entertainment, assisted diagnosis/surgery, robotics, etc.

One of the most used techniques to obtain 3D information from a scene is triangulation, and more concretely, triangulation-based laser scanners. Since their formal appearance in scientific publications, in 1971 [SS71], there have been contributions for solving inherent problems like occlusion avoidance, accuracy improvement, acquisition speed, shape description, etc. All of the methods for obtaining 3D points of a scene is accompained with a calibration procedure, and this procedure plays a decisive role in the performance of the acquisition device.

The goal of this thesis is to provide a holistic approach to the problem of shape acquisition, giving a wide survey of triangulation laser scanners, testing the performance of different systems, and to give contributions for both improving acquisition accuracy under adverse conditions and solving the calibration problem. In addition, the calibration approach is based on previous works that used projective geometry to this end.

Chapter 1

Introduction

Perception is the interpretation of specific sensations from the visual, auditory, somatosensory and olfactory sensation systems. Vision is the most important physical sense of human beings. It gives us the ability of seeing, that is, to have perception of colour, texture, shape and depth. The goal of this chapter is to give an overview of how this physical sense works, starting with reviewing how the image is formed in the eye, following the visual flux up to superior cognitive layers, until the perception of shape takes place.

1.1   Introduction to 3D metrology

One of the goals of cognitive scientists is to understand how the human brain converts the raw data from each of the sensing systems into a coherent and meaningful interpretation of the outside world. Figure 1.3 shows a side section of the human eye, where the parts contributing to the image formation have been labelled. As can be seen, there are numerous similarities with a video camera, since we can find the cornea, which is a transparent coating, the sclera, which is an outer cover composed of a fibrous coat that surrounds the choroid, the pupil, which sets the iris aperture for controlling the amount of light impinging on the retina, which plays the role of the image sensor, and finally the crystalline, which plays the role of a lens to focus the image on the retina.

The perception that humans have of shape and depth is something that falls beyond the simple eye image *acquisition* process. There are numerous scientific studies which relate the activity of certain cortical areas of the brain with the binocular disparity [BBvEC99, Bac00, BFPH01], which is the basis for shape perception. Figure 1.4 gives an overview of the streams of processing in the primate vision system and shows how the anatomical connections

can be viewed as a feed-forward hierarchy in which each neural layer extends the processing carried out by layers below it [DJMU85]. Traditionally, the computer vision and artificial intelligence communities have been trying to imitate the human ability of describing our surroundings using computer algorithms. The topics covered by computer vision are diverse and stratified in levels of complexity, i.e. high level, like object recognition and feature extraction, mid level, like 3D perception, geometric vision, and low level, like point or edge detection. There is still an analogy with the human vision system, which involve higher levels of cognitivity as the *output* of the vision system increases in complexity.

If we focus our attention to 3D or shape perception, it is obvious that a stereo camera system keeps a strong analogy with a human head. A stereo camera system is able to extract the 3D coordinates of a point in space by computing the intersection of the two lines that join the point and the two focal points of both cameras. A camera has a rectangular, planar image sensor which plays the role of a retina, which is often called the *retinal plane* within the computer vision community. However, even when the similarities with eyes are significant, there is some important differences between video cameras and eyes. These differences lay mainly in the fact that the biological retina is not a flat sensor, but it is concave. In addition, biological retinas' composition is not homogeneous neither in the distribution nor in the types of *light transducers*. On the contrary, they are spread inhomogeneously such that there is more concentration in the centre of the retina than far from it in the radial direction. Figure 1.1 shows the distribution of light transducers on the retina. There can be observed that a blind spot exists exactly where the optical nerve collects the visual information. Moreover, there are different types of transducers, namely cones and rods, which make the central region of the retina, good at perceiving colour, while the outer area perceive better the shapes or textures, insensitively to the colour, and the most lateral areas perceive only movement. Figure 1.2 show an enlarged view of the cones and rods present on the retina. They are, in essence, very specialised bipolar neurons which have developed special structural features to carry on their task, where rods are more sensitive to light than cones, but do not respond to colour impulses, and cones are much less sensitive to light than rods, but they are very good at discriminating colours. All these features lead the human eye to be conceived as a very directional visual sensor, and it has been proved that the orientation and, of course, the separation of the eyes play a very important role in the perception of depth or shape, which are less important in camera-based stereo systems. However, important efforts are being spent in order to mimic the human vision system *as is*, that is, in developing *foveal* vision systems. Their current use is mainly intended to study attention and
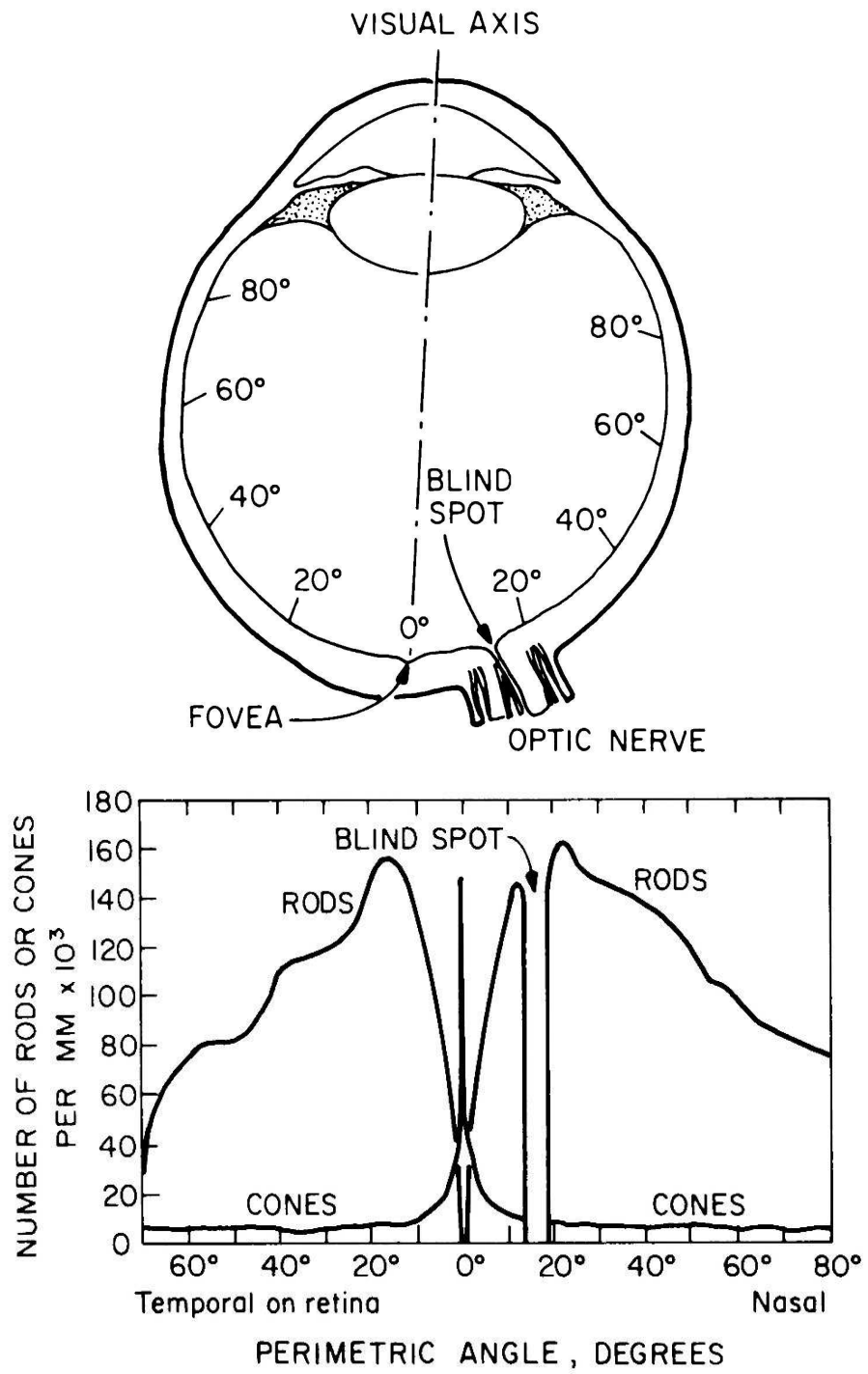
Figure 1.1: Concentration of light transducers on the retina.
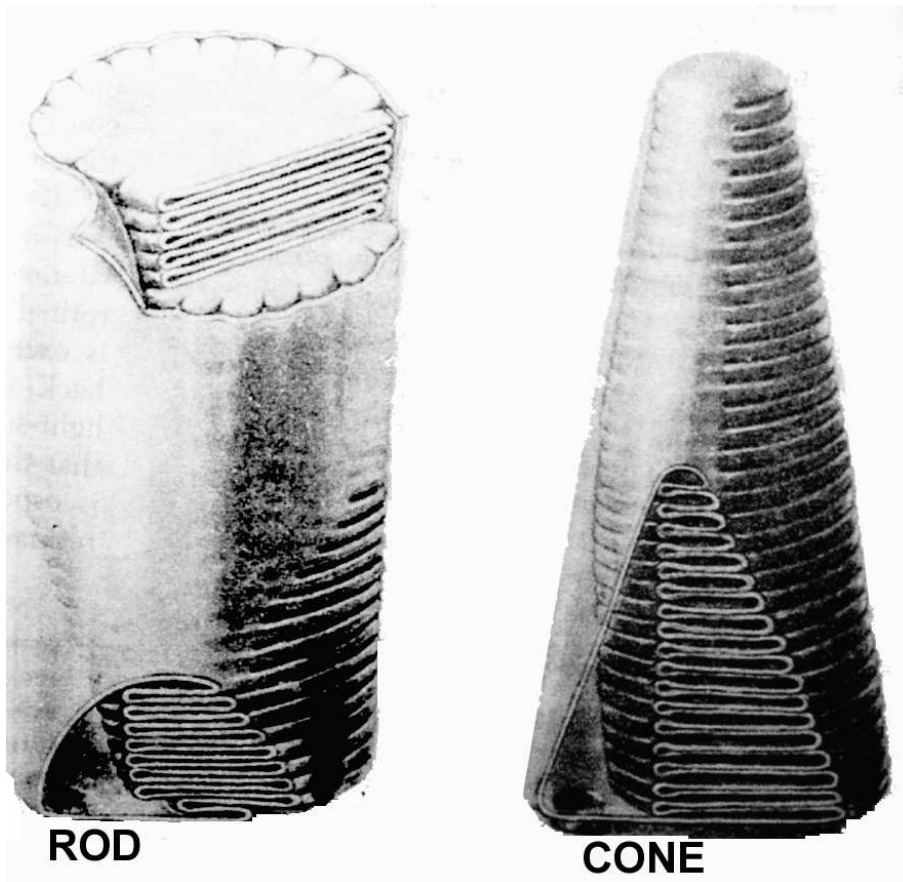
Figure 1.2: Cones and Rods present on the retina.
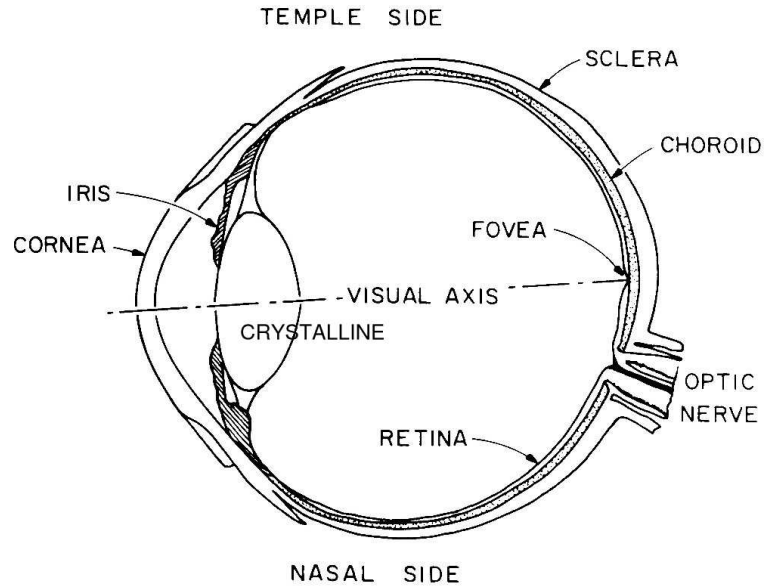
Figure 1.3: Side section of a human eye.

for object tracking purposes. The current foveal sensors are either CCD or CMOS, built with maximum image resolution at the centre, decreasing more and more toward the periphery [LCM00, XM02, XM03]. Figure 1.5 shows an example of how these sensors perceive the scene.

Although we, humans, are very good at perceiving shape, depth, colour and texture, due to thousands of years of evolution and specialisation, the recent trends of mankind life, require new skills for making accurate measurements, leading the concept of metrology to be more important than ever. The term comes from the Greek *metros* - measure, and *logos* - treaty, hence we can state that metrology is the science of measuring, and that to measure is to compare with something (a unit), and this unit is taken to be the basis for comparison. In particular, 3D metrology is a very interesting technology and is being applied in diverse fields like industry (rapid prototyping, reverse engineering, modelling, etc.), 3D medical imaging, fluid analysis, human body measurements, quality assessment, robotics, entertainment, etc... There has been a great upsurge of such devices since the former appearance of the first Coordinate Measuring Machine (CMM), which are contact measuring devices capable of measuring points with a very high accuracy (up to a few units of micron). Besides, this process is extremely slow, it must be carried out by a skilled operator and it is a very expensive technology, and hence, it is limited to high end industrial processes. On the contrary, non-contact measuring
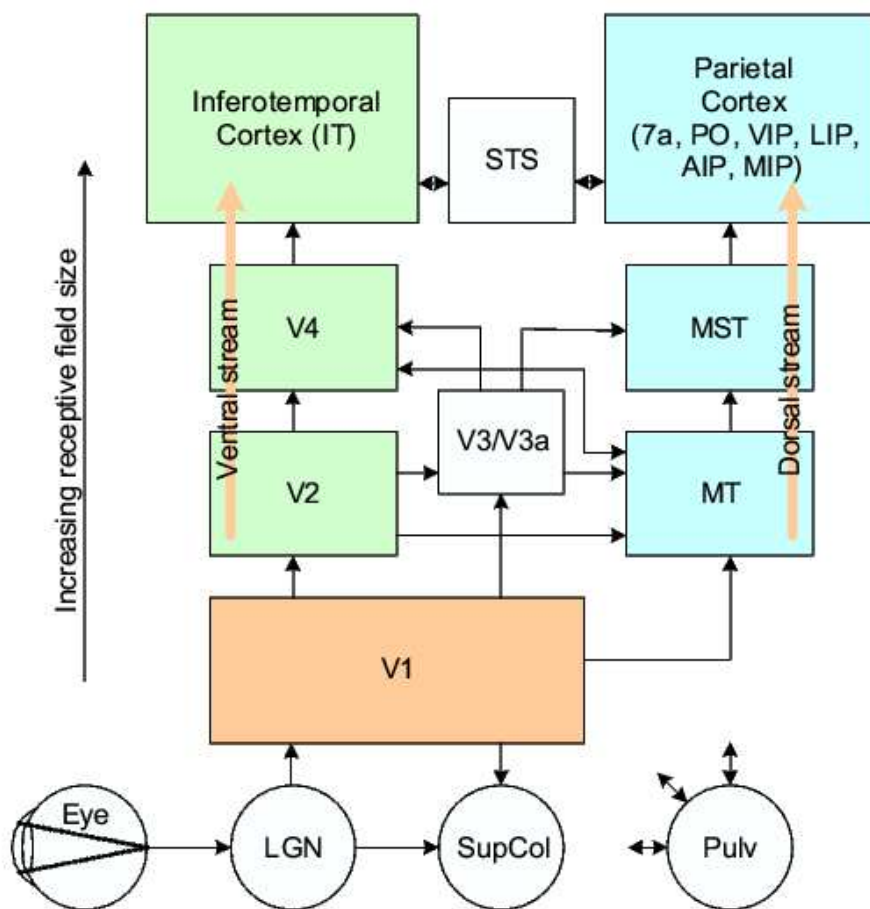
Figure 1.4: Interconnections in the vision system. Sub-cortical areas: LGN, Lateral Geniculate Nucleus; SupCol, Superior Colliculus; Pulv, Pulvinar nucleus. Cortical areas: MT, Middle Temporal; MST, Middle Superior Temporal; PO, Parieto-occipital; VIP, LIP, MIP, AIP, Ventral, Lateral, Medial and Anterior intraparietal respectively; STS, Superior Temporal Sulcus.

Figure 1.5: A text as seen with a foveal vision sensor.
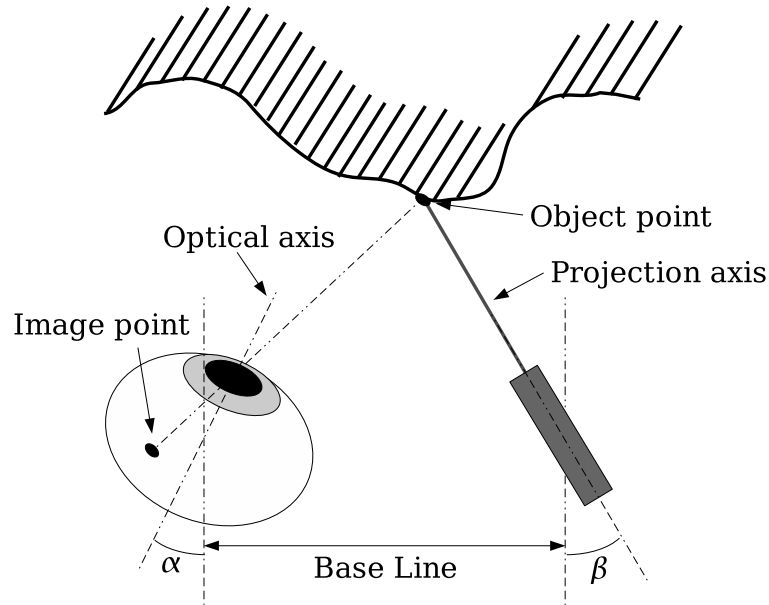
Figure 1.6: The basic 3D laser scanner.

devices, the so called 3D scanners, are being more and more present in many fields, and yet they are getting into the home environment in the form of low-cost 3D cameras.

One of the approaches to 3D sensing by non-contact devices is the widely spread active 3D scanner devices. They are called active, because they change the lighting of the scene by projecting structured light on it. The way they obtain 3D information is by using a camera for measuring the disparity between the projected structured light and the image of the scene. Among the structured light devices, we focus our attention to the case of triangulation-based 3D laser scanners. Figure 1.6 shows a top view of the basic diagram which builds up such system. Note that there appear important concepts and quantities like the base line, the optical axis, the laser axis, the angles α and β with respect to the vertical, quantifying the orientation of the *eye* and laser, etc. These quantities model the geometric aspects of the system. The eye appearing in that figure is obviously replaced by a camera in real applications. Note that, just like in the case of the human vision system, the position and orientation of the eye/camera is important for shape/depth sensing, as well as the orientation and position of the laser projection. In addition, a laser stripe or spot can be narrowly focused even at great distances, hence providing even more directionality than the foveal vision.

1.2   Context and motivation

3D scanners are today's most accurate non-contact measuring tools. However, there are two important issues that can be improved. The former concerns the identification of the geometrical parameters of the system, since they allow the equations to be identified and solved in order to provide the 3D points of reconstruction. The process for obtaining these values is called calibration, and it is solved in a variety of ways throughout the literature. In particular, our research group has developed works toward the calibration of both structured light and stereo rig devices [Sal97, BMS98, SAB02, FS02, SPB04]. In addition, contributions toward the guidance of both mobile and underwater robots have been undertaken within the group, specially applying structured light for short range computation and the fundamental matrix estimation for mobile robot motion measurement [RSB98, AS03]. The contribution of this thesis to this end, focuses on the projective geometry approach. The second problem concerns the detection of the laser stripe. It is worth regarding that the geometric figures that are reconstructed are points, and points are zero-dimensional, that is, they do not have volume nor surface. It is evident that the size of a laser spot is, by no means, an infinitely small point. Instead, it is an electromagnetic radiation that can be focused down to about 2μm using the current lens manufacturing technology. Hence, it seems very sensitive to provide a means of detecting a single point as much representative of the laser spot as possible. The light energy distribution of the stripe follows a Gaussian pattern, hence it is assumed that a good candidate for this *representative point* is the light peak of the Gaussian pattern. For this reason, the techniques or algorithms for obtaining the pixel of the image that best quantify this point are called numerical peak detectors. The need for a projective geometry-based calibration procedure is clear, and several authors have contributions to this end. With this approach, it is not necessary to use accurate mechanical positioning parts for identifying the laser plane/line equation, and this feature is very interesting when the calibration in hazardous environments such as underwater applications is needed.

1.3   Objectives

This thesis has been developed with the aim of solving the issues of calibration using projective geometry and the obtention of peak position estimation under extreme conditions. With these goals in mind, the steps which led through the development of the thesis can be enumerated as follows:

- Applications such as human tissue scanning or very fast range scanning motivated the study of a new technique for detecting the peak of a laser stripe. The reason is that projecting laser light on a piece of human tissue at a microscopic scale, a deep penetration of light inside the surface is observed, as well as a very significant diffusion of light through the tissue. This effect is similar when scanning a translucid material at such scale. In addition, applications such as those requiring a very high speed scanning, keeping the laser power at a visual safe level, are subject to a low signal-to-noise ratio when imaging the stripe with the camera. This is materialised as a very low light level on the stripe image, and hence, hardly distinguished from the background. In both situations, noise is the limiting factor for obtaining a good peak estimation. The former, because the noise level is very high and distributed over a wide area of the surface, and the latter, because the laser stripe light level is very close to the background noise level. The proposed peak detector has shown to perform very well under both conditions.

- There exist many calibration algorithms for triangulation-based laser scanners, all of them with their pros and conts in particular aspects. Specially interesting are those which involve the modelling of the laser-camera relationship using projective geometry, since they avoid the identification of physical parameters which describe the position and orientation of the camera and laser to each other. The only condition is to assume a linear model for the camera, that is, a pinhole model, and this assumption is valid as long as the optics are chosen so that no image distortion is present. In addition, in the presence of image distortion, if the intrinsic parameters of the camera are modelled through calibration, the projective model is still valid since distortion can be rejected. However, the existing projective calibration approaches suffer from the fact that either the laser plane can not change its pose with respect to the camera, constraining its flexibility, or a limited number of plane poses are used, giving a low scanning resolution. In this thesis, a new method is proposed such that only a few plane poses are used for calibration purposes, but arbitrarily high scanning accuracy is obtained. In addition, this method allows for the laser plane to change its orientation with respect to the camera.

- The calibration method is based on the modelling of the geometrical relationship between the laser plane and the camera image plane, by means of a projective transformation. This transformation is obtained

from point to point correspondences, and these point correspondences have to be obtained as accurately as possible, and they must correspond to both the 3D coordinates and the 2D image of the laser stripe. To this end, the cross-ratio of four collinear points is used, so that it is computed on the image, using four measured points in pixels, and extrapolated to the 3D coordinates in order to obtain the fourth 3D point, collinear with other three well known points. The complete quadrangle is a geometric figure which is defined by a slanted rectangle, so that no two edges of it are parallel. If the points at the four corners are known, three other points are automatically defined by the intersections of the lines joining the other four points. This fact allows to keep three 3D points known, using the cross-ratio to get the coordinates of the fourth, and hence obtaining accurate point to point correspondences.

1.4   Thesis outline

This thesis is structured in 7 chapters and 2 appendices, following a natural stream of what is the problem of 3D shape acquisition using laser scanners. Chapter 2 reviews the triangulation-based laser scanners since their former appearance in 1971, proposing a classification of such devices according to their principle of operation, and summarises this classification in the form of an indexing table. More than 120 articles have been reviewed, including references of the latest conferences. One of the topics covered by the survey, the projective approach, has been taken as a starting point for developing a new calibration method.

Chapter 3 describes a new technique based on the use of digital filters for obtaining a reliable peak position even at very low signal-to-noise ratios. The proposed peak detector is based on a well known result of the convolution properties, and the application of FIR filters for obtaining accurate measurements. The choice of the filter coefficients is based on the choice of the cut-off frequency and the transition band abruptity. This choice has been made empirically in our lab experiments, although an automated method would be helpful. However, the development of such method falls beyond the scope of this work.

Chapter 4 proposes a novel calibration strategy based on the application of projective geometry, which includes a method for obtaining the best 3D point correspondence, given its 2D point image. This method is based on the application of the cross-ratio and the use of a well known geometry in the shape of a complete quadrangle. This quadrangle is used as a calibration target and is useful due to its well defined 7 points, of which, only 6 are

used. The method allows for the reconstruction of objects with arbitrarily high scanning resolution, even when only a few equally-spaced scan positions are needed for obtaining the calibration parameters.

Chapter 5 exposes the experimental results which have been undertaken in our lab facilities. Chapter 6 discusses the conclusions of each contribution and gives directions on how the results could be improved and the further work which can be carried out from this work. Appendix A describes several applications that were carried out as a previous work which led us to the development of the thesis contributions. Finally, appendix B reports on the notation used throughout this work.

Chapter 2

Triangulation 3D laser scanners

Non-contact three-dimensional digitisers are becoming more and more used tools in engineering, in substitution of contact Coordinate Measuring Machines (CMM). Non-contact methods are faster, do not change physical object properties, and are cheaper. Among the different methods available (stripe pattern, coded structured light, rainbow, etc...), laser scanning is one of the earliest ones, even increasing its use today. The calibration of a three-dimensional digitiser is a very important issue to take into consideration if good quality, reliability, accuracy and high repeatability are the features which such a good digitiser is expected to have. The purpose of this chapter is to report an exhaustive survey on non-contact triangulation-based laser scanners, making special attention on the calibration methods employed.

2.1 Introduction

A three-dimensional digitiser is often called range finder too, stating its capability for acquiring range or depth information. These devices use to grab "range images" or "$2\frac{1}{2}$D images", which are dense arrays of values related to the distance of the scene to a known point or plane.

Many laser scanning three-dimensional digitiser systems use one or more standard CCD cameras and scanning laser slits, and triangulation techniques are widely used in the 3D reconstruction of the scenes. Some of these systems operate grabbing one image per each sweep step, so it is a time consuming process, yielding to a big amount of images. However, using high speed vision modules, rates of up to 1.6 complete scans per second have been achieved, as is the case of the *MINOLTA vivid 700*. Besides, speeds of up to 3 range images per second have been achieved using standard CCD cameras. The operation of such systems (for example CubicScope) rely on

the fast switching of the light source and the CCD shutter operation in or-
der to get space-time encoded images. A lot of manufacturers of scanning
three-dimensional digitisers exist: *Intecu*, *Cubiscope*, *3DScanners*, *Philips In-
dustrial Vision*, *DigiBotics*, *Steinbichler Optotechnik*, and more. A complete
list of manufacturers with links to their own web pages can be found in
http://www.3dlinks.com/hardware_scanners.cfm.
In real time applications range information is required continuously, without
gaps in the data. This limits dramatically the use of commercial three-
dimensional digitisers in such fields.
Research has been carried out by [GKC90, YSYI94] toward the development
of *Smart sensors* capable of delivering up to 1000 $2\frac{1}{2}$D images per second,
but only laboratory prototypes were developed.
This chapter is intended to give an overview of laser slit scanning three dimen-
sional digitiser systems paying a special attention to their calibration meth-
ods, reporting a number of references for the interested reader. In addition,
the different surveyed authors and their corresponding systems and methods
are classified in table 2.1 in order to provide the reader with a reference tool.
In section 2.3 an illustrative classification table is proposed, while section
2.4 expands the proposed classification presenting an in-depth description of
some of the systems. There are numerous surveys reporting the different 3D
scanner technologies, but none of them is focussed on triangulation-based de-
vices exclusively. The state-of-the-art review has shown that there is a great
variety, even when it is a very specific topic. In addition, this survey evi-
dences the evolution that this area has experienced, not only in technology,
but also in the approaches that improve both calibration and reconstruction.

2.2   The principle of triangulation

This section introduces the principle of triangulation focusing its use in laser
slit scanning. The principle of triangulation is always accompained by the so
called problem of correspondence, which is still unsolved in many cases, and
configures the stumbling stone in complex problems like robot navigation or
simultaneous localisation and mapping. The reader is pointed to Hartley
[Har97] for an optimal global solution on the triangulation problem. The
typical situation in which such principle is applied is shown in figure 2.1
where a point P in 3D space is *seen* by two different cameras. If the spatial
relationship between both cameras is accurately known, the position of P
can be found by its two projections (p' and p'') on the corresponding image
planes, computing the intersection of lines $P - p'$ and $P - p''$.
The spatial relationship between both cameras can be easily found by apply-
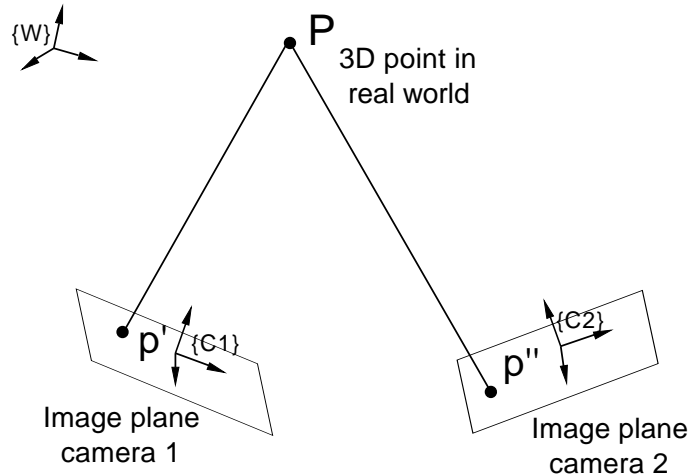
Figure 2.1: Triangulation principle using 2 cameras.

ing a standard camera calibration procedure to be chosen among the multiple existing ones, depending on how distorted are the captured images. Different camera calibration methods with and without distortion consideration can be found in [HJTS82, Fau93, Tsa87] and[WCH92]. In addition a survey has been published recently by Salvi et.al. [SAB02] comparing the performance of different methods in terms of their accuracy.

But, unfortunately the biggest problem here is: how can P be uniquely identified in both image planes ? i.e. how sure can be stated that both p' and p'' correspond to the projection of P ? The answer to these questions is not unique and a lot of research is being done in this direction. This problem is widely known as *the correspondence problem*, which is very difficult to solve robustly in multiple camera systems, unless very well defined corners are imaged. Besides, one of the techniques that works very well for identifying points in 3D space is the use of Active Vision. The term *active* means that some well known light pattern (typically laser or coded/uncoded structured light) in different geometries (points, planes, crosses, parallel bars, etc.) is projected onto the scene, so an illuminated point in the 3D space is well seen by both cameras, and hence, the 3D coordinates of P are easily computed by triangulation. An excellent survey on active range finders was published by Besl [Bes88]. Moreover, a thorough survey comparing the use of different coded structured light schemes to reduce the correspondence problem has been published by Batlle et. al. [BMS98].

In addition, if the light source geometry, that is, its pose and scanning movement, is well known, one of the cameras can be replaced by the light
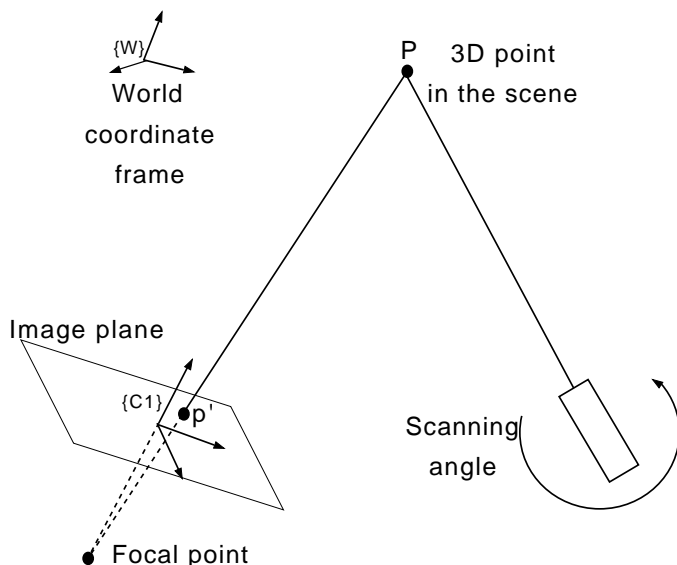
Figure 2.2: Laser scanner system with scanning light source.

source and 3D information can be obtained by triangulation as well. This procedure is widely used by commercial and research prototype range finders. Figure 2.2 shows how camera 2 in figure 2.1 has been replaced by a sweeping laser light source. Using this technique the camera model has to be computed using a calibration procedure, and the line or plane laser equation must be known in order to compute its intersection with the line $P - p'$.

In order to estimate the laser equation, the sweeping angle must be measured by means of some external mechanical or optical rotation sensor (potentiometers or encoders).

Equations 2.1 show how a 3D point can be computed, knowing one point of the sweeping laser plane (P_0'), the $P - p'$ line equation (with direction vector $\overrightarrow{v}$) and the plane's normal vector $\overrightarrow{n}$ (which depends exclusively on the sweeping angle). Figure 2.3 shows the laser plane and camera arrangement where these equations can be applied. Till now, the triangulation problem has been regarded as the computation of the intersection point between two lines or between a plane and a line. However, even when in the ideal case this computation is straightforward, the presence of noise turns it into a minimisation problem. The case of two view triangulation is well documented and discussed by Hartley [Har97], and it consists of computing the point closest to the two lines or rays, according to some minimisation criteria. In the case of laser scanners, a good triangulation reconstruction is up to the accuracy

Sweeping
Angle

$\vec{n}$

Laser
plane

P

P'o

$\vec{v}$
p'
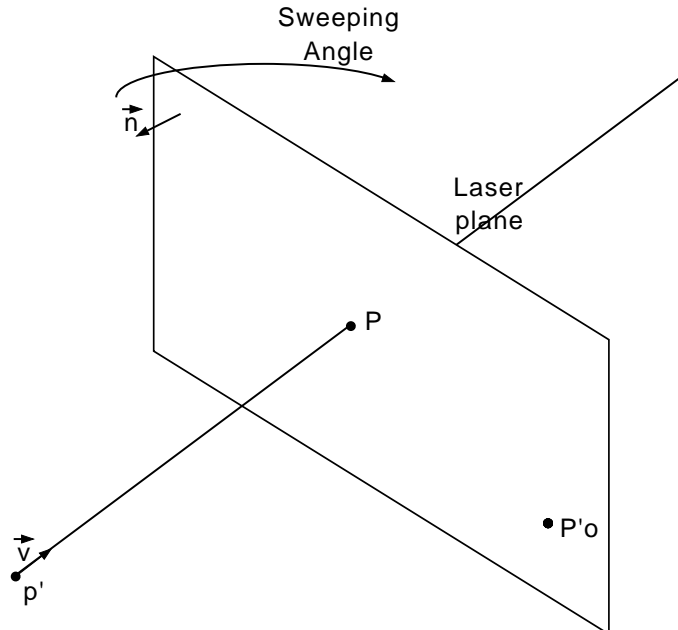
Figure 2.3: 3D reconstruction by triangulation using a laser plane and a camera.

with which the laser light peak is located on the image. This peak gives the position of point p' in figure 2.2, hence the need of using the best peak detector if the 3D reconstruction accuracy has to be maximised. Chapter 3 describes a novel algorithm for peak detection, and compares its performance with 5 other existing methods.

$$\begin{cases} P = P_0 + \lambda \vec{v} \\ (P - P'_0) \cdot \vec{n} = 0 \end{cases} \tag{2.1}$$

2.3 Proposed classification

Table 2.1 shows the reviewed three-dimensional digitisers, reporting the main features they exhibit, and showing a structure based on the calibration methods that yield to the triangulation reconstruction. For each of them, several authors are referenced, reporting the main features of the systems they have proposed, and including an illustrative sketch. This table is intended as a fast reference for the readers. The classification is structured in the following categories:

- **Pose**: The estimation of the laser plane/line equations is performed by

using accurate mechanical positioning stages. The calibration of the whole system may include a former camera calibration.

- **Projective**: The system is calibrated as a whole, and involves the identification of the projective transformation between the laser plane and the image plane.

- **Euclidean**: The calibration involves euclidean geometry measurements, but they are performed either automatically, using a former camera calibration and geometric constraints like collinearity or coplanarity, or by manufacturing a well-known object which is used as a reference, or calibration pattern.

- **Time Mux**: Involves a camera calibration process, but the relationship between the image plane and the laser plane is identified by coding the pixels which belong to each laser plane position. This coding is achieved by switching the laser light source.

- **LUT**: A Lookup Table is generated in order to state the relationship between the scanning position and known depths.

Table 2.1: Classification of three-dimensional digitiser systems.
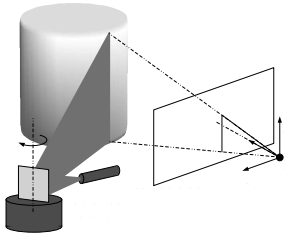
Type	Characteristic	Features
Pose		*Auth:*Shirai et al.[SS71], Agin et al.[AB73], Popplestone[PBAA75] *Calib:* [SS71, PBAA75] Accurate pose measurement of system devices [AB73] (A)-Camera calibration (B)-Estimation of laser slit position by image measurements *Features:* • Single slit • Rotating mirror reflects only the laser slit

Table 2.1: (continued)
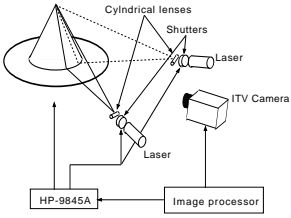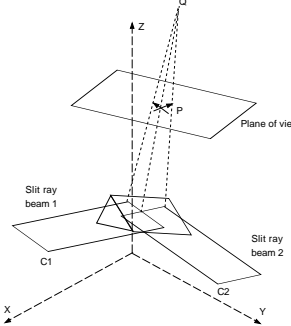
Type	Characteristic	Features
Pose		*Auth:* Sato[SKF82] *Calib:* Point of light attached to the turntable at known radius and heights. *Features:* • Two slit projection • Turntable operation • Static laser and camera set • Minimisation of occlusions
		Auth: Nakano[NWK88] *Calib:* Accurate pose measurement of system devices *Features:* • Two slit projection • Laser angular scanning • Eliminates reconstruction errors in specular surfaces

Table 2.1: (continued)
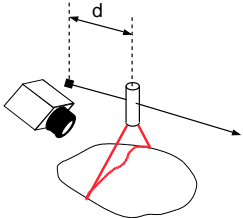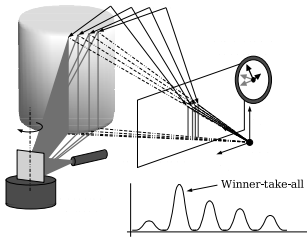
Type	Characteristic	Features
Pose		*Auth:*Champleboux[CLSC92][CLSB92], Davis[DC01] *Calib:* Champ: (A) - Camera calibration using NPBS method. (B) - Accurate mechanical measurement and $R^3 \rightarrow R^3$ transformation identification Davis: (A) - Camera calibration using Tsai method. *Features:* • Single slit • Linear scanning • *Davis*: Single camera stereo (wedge mirror)
		Auth: Brajovic et al. [BMJ01] *Calib:* Acurate pose measurement. *Features:* • Single slit • Rotating mirror • One-per-row slit peak detectors (winner-take-all) [LRMM88] • Row-parallel architecture

Table 2.1: (continued)
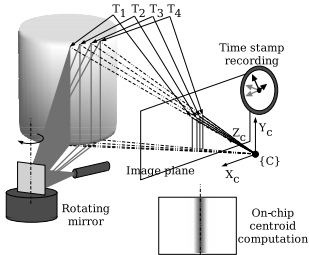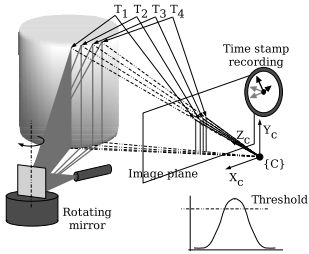
Type	Characteristic	Features
Pose		*Auth:* Oike et al.. [OIA03a][OIA03b] *Calib:* Acurate pose measurement. *Features:* • Single slit • Rotating mirror • One-per-row slit sub-pixel centriod computation • Row-parallel architecture • Up to 260Kframes per second
		Auth: Gruss et al.[GKC90][GTK92], Kanade et al.[KGC91], Hori et al.[HBK94], Baba et al.[BKK97], Araki et al.[ASP87] *Calib:* [GKC90][GTK92][KGC91] Line-of-sight identification using a specially designed calibration target. The target is accurately positioned using a mechanical system *Features:* • Single slit • Rotating mirror • Cell-parallel architecture • Accuracy is determined by slit detection method • Slit detection by threshold of the signal level

Table 2.1: (continued)
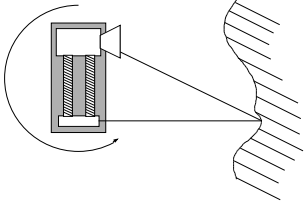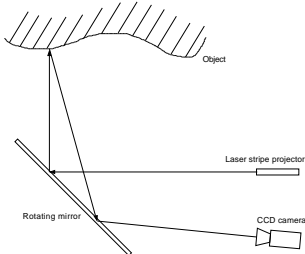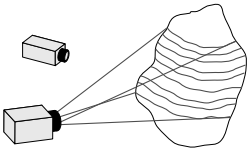
Type	Characteristic	Features
Projective		*Auth:*Chen & Kak [CK87] *Calib:*Projective geometry *Features:* • Single slit • Fixed geometry, rotates the whole system
		*Auth:*Reid [Rei96] *Calib:*Projective geometry *Features:* • Single slit • Fixed geometry • Rotating mirror reflects both the image and laser slit
		*Auth:*Huynh [Huy97], Park [PDK01] *Calib:*Projective geometry *Features:* • Multiple slit projection • Fixed geometry • No mechanics for slit scanning

Table 2.1: (continued)
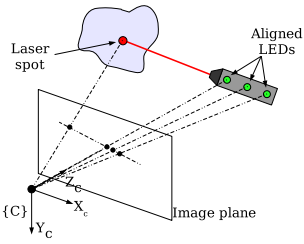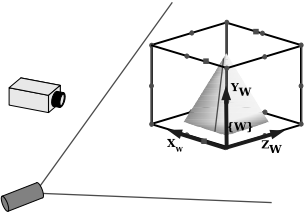
Type	Characteristic	Features
Euclidean		*Auth:*Takatsuka[TWVC99], Furukawa[FK03] *Calib:* (A) - Camera calibration (Pinhole). (B) - Line/Plane identification *Features:* • Single slit/spot • LED-marked pointer • No scanning mechanics • *Wand* always visible
		*Auth:*Chu[CHJ01] *Calib:* Specially designed, well-known target. *Features:* • Single slit • Known cube always visible • No scanning mechanics • Reconstruction by linear relationships

Table 2.1: (continued)
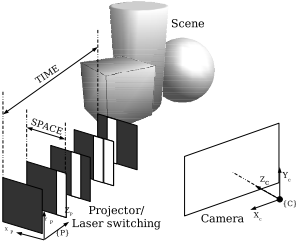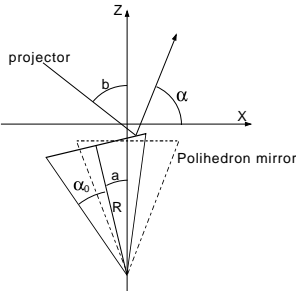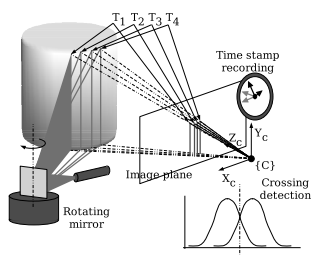
Type	Characteristic	Features
Time Multiplexing		*Auth:*Sato[HS95][Sat95][HS96] *Calib:* (A) - Camera calibration (Tsai [Tsa86]). (B) - Plane equation identification according to the space coding. *Features:* • Single slit • Rotating mirror • Compromise between acquisition speed and accuracy • Use of gray code in the pattern projection • Pattern shifting for accuracy improvement
		*Auth:*Yu[YZWQ96][YZWL98] *Calib:* Square pattern projection. System layout estimation using geometry analysis *Features:* • Single slit • Rotating mirror • Baseline error compensation • No special code spedified, although the system is versatile enough for adopting any codification • Tradeoff between acquisition speed and accuracy

Table 2.1: (continued)

Type	Characteristic	Features
LUT		*Auth:* K.Sato et al.[SYI94a], Yokohama et al.[YSYI94] *Calib:* Lookup table generation by finding the relationship between the scanning *time* and known depths. A plane is used as a calibration target *Features:* • Single slit • Rotating mirror • Cell-parallel architecture • Slit detection using comparison between two photosensitive areas in each cell

2.4 Systems and methods for shape acquisition

Shape measurement is intended to be as much accurate as possible. The calibration procedure together with the stripe detection method play an important role in this sense. At the end, accuracy is dependent on both the accurate measurement of the relative pose between the imager and the light projector, and in how certain is the stripe detection in the image itself. The imager and slit projector pose may be measured using very accurate mechanical means, although it is not applicable to the systems which are subject to a modification of their *default* mechanical arrangement.

In addition, sub-pixel accuracy stripe detection methods may be applied in order to minimise the uncertainty. A comparison between five different such methods can be found in [TFFN98], and a new contribution on the obtention of the accurate peak location is reported in [FSCP04]. This section discusses the methods surveyed in section 2.3, addressing the general description of each system according to the dissertation of the authors. It follows the same structure as table 2.1. A special attention is made on calibration and recon-

struction issues.

2.4.1   Pose measurement

Shirai & Suwa [SS71] and Agin & Binford [AB73] used a video camera together with a plane-shaped laser light source projecting a slit onto a scene. The slit scanned the scene using a mirror attached to a stepper motor, and the camera was used for the acquisition of the stripe image. Shirai & Suwa calibrated their system by measuring the relative pose between the camera and the projector. The use of this method is strongly discouraged if an accurate mechanical positioning system is not available. Agin & Binford used a calibration procedure consisting on a former camera calibration, followed by the estimation of the light projector position by grabbing a series of stripe images at known angular positions of the projector. A relaxation algorithm was used in order to minimise the position error. The accuracy in scene reconstruction is mostly dependent on the right choice of the camera calibration method.

Popplestone [PBAA75] used a range finder that relied on the operation of a *vidicon* TV camera, measuring the time that a TV line scan takes until it achieves the image of the laser slit. The apparatus returned a number of clock pulses per each TV line, using a fast clock signal.

Sato et al. [SKF82] described a system consisting on two laser slit projectors and a CCD camera, but the object was placed on a turntable, being the laser-camera set static. The two laser sources were used in order to minimise the occlusions. A light target together with a known object was used in order to calibrate the system. A point of light was attached to the turntable at known radius and heights and the turntable was operated, grabbing several images of it. The accurate measurement of the laser projector location was finally used in order to obtain three-dimensional data by triangulation. The reconstruction accuracy is dependent on how accurate is the location of the light peak on the target.

Champleboux et al. [CLSC92] and [CLSB92] described a range imaging sensor where the laser was linearly displaced using a linear actuator. It was used together with an object localisation algorithm for the matching of human face surfaces. An accurate positioning mechanical system was used in order to identify the light plane equation. Then, the system calibration was done by identifying the transformation $G : R^3 \rightarrow R^3$ defined by the relation (X,Y,Z)=G(u,v,d), where (u,v) are the pixel coordinates of point (X,Y,Z) and d is the linear displacement of the laser slit. Davis [DC01] developed a very low-cost scanner based on the conventional stereo approach. Figure 2.4 shows the details. Instead of two cameras, a wedge-shaped mirror is used
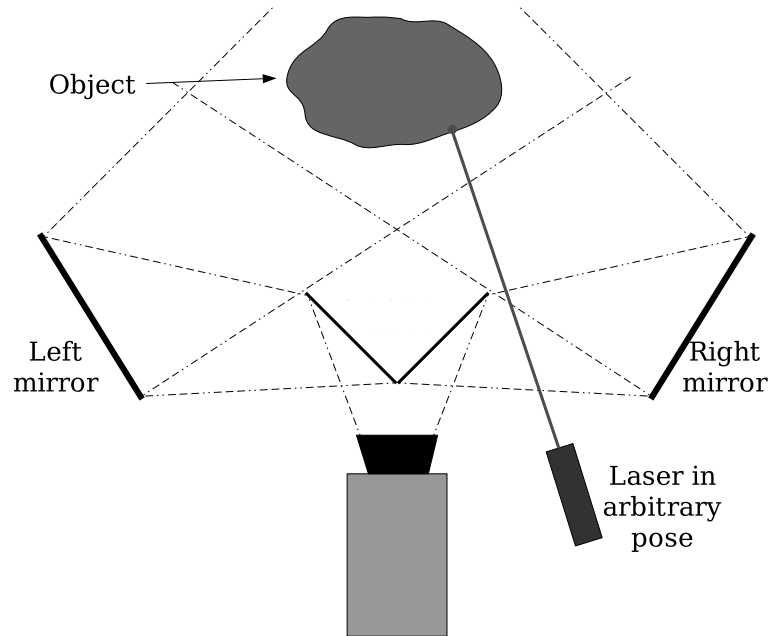
Figure 2.4: A stereo rig with only one camera and a laser stripe.

for generating two side images on a single camera. The camera is calibrated using the method of Tsai [Tsa86], and the 3D reconstruction is achieved by computing the intersection of the two optical rays passing through the same 3D point. In this work the laser equation is not computed, since it is used only for solving the corresponding problem. The system performs acceptably good, considering its low cost. However, even when the issue of syncronisation between the cameras in a common stereo rig is obviously solved, the sync between the camera shutter and the laser movement should be considered in order to minimise even more the inaccuracies. Oike et. al.[OIA03a] developed a smart image sensor for real-time and high resolution 3D measurements. The features of this device allow high speed measurement and incorporates circuitry for centroid-based peak detection and is able to achieve very low noise 3D information. Figure 2.5 shows the strategy for the sub-pixel accuracy centroid computation. The sensor is able to detect both the position and the intensity profile of the projected light. The frame-rates are 260K fps for 128x128 pixels and 32.6K fps for a 1024x1024 pixel frame. In addition, Oike et. al. [OIA03b] developed another similar device intended for high speed position detection. It includes a parallel row architecture, which allows for a considerably higher frame-rate. The edge of the illuminated pixels is quickly detected and provides enough information for triangulation.

| 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

a) Image Acquisition

| 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

b) Left edge found

| 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

c) Right edge found

| 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

d) Sub-pixel accuracy Centroid Computation
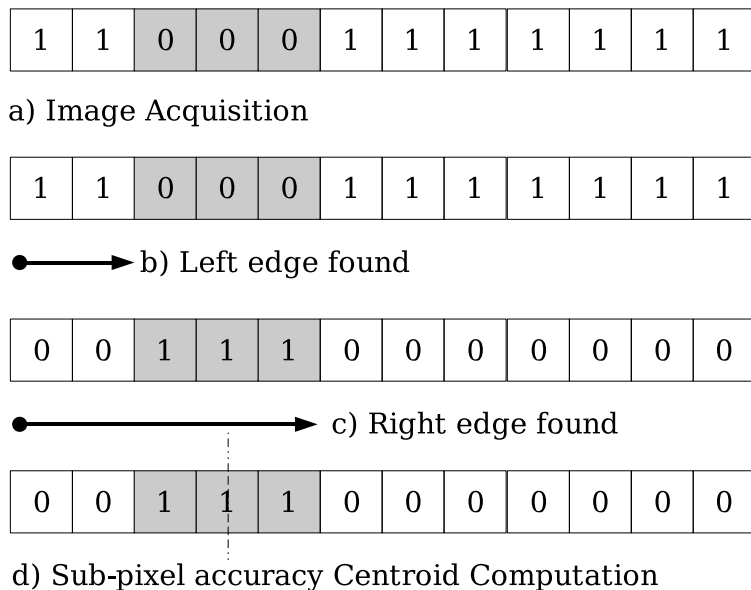
Figure 2.5: Centroid Computation Strategy in Oike et.al.

The peak detection circuitry allows a resolution of up to $1/256^{th}$ of a pixel. Araki et al. [ASP87] proposed a new method for measuring the *time* at which the slit image projects onto a photosensitive cell of a specifically designed discrete array. Each of the photosensitive cells has a timer register associated which is incremented until the image of the laser slit illuminates its associated photosensitive cell. This kind of operating procedure allows a very fast scanning, since each of the photosensitive cells is concurrently *processing* by its own. Hence, this operating procedure is called **cell-parallel** processing. Two optical switches are used in order to indicate the start and the end of a scan period.

A simple threshold circuit was used in each cell in order to detect the light stripe image. The output of this circuit was used to stop the timer counting and store its value on the register itself. Once the light stripe mechanism fulfils a complete scan period, the register array is configured such that a big shift register is formed. Then the time values are transferred serially to the host computer.

A rotating mirror arrangement similar to that of [HS96] or [YZWQ96] is used in this implementation.

Gruss and Kanade [GKC90, GTK92, KGC91] proposed a VLSI version of a similar system. The conceptual difference with the system by [ASP87] is that the time computation is made analogically, using capacitor-based timing cir-

cuits, and a periodic sawtooth signal. The stripe detection was accomplished by using simple threshold circuits integrated in each cell.

Analog time stamp recording is used instead of digital timers in order to avoid EMI from the switching clock signals to affect the photosensitive cells output.

Kanade et al. [KGC91] and [GTK92] described a prototype implementation of the system in [GKC90].

In [GTK92] the calibration target is a planar surface out of which a triangular section has been removed. This target is mounted on an accurate 3DOF[1] positioning device so its pose is accurately known. The calibration procedure is as follows: 1) the line-of-sight rays for a few cells are measured, and 2) a pinhole-camera model is fit to the measured line-of-sight straights in order to approximate the line-of-sight equation for all sensing elements.

The target is positioned so that its surface is parallel to the reference xy plane. Then it is moved along the search path (see figure 2.16) until an occlusion (the photosensitive cell does not detect the slit) occurs. Finally, the 3D position of the target is recorded, and the process is repeated until both the bottom and top edges of the triangle are found. This procedure is made for different depths, so a set of *triangles* is recorded. The right corner of the triangle is found by calculating the bottom and top lines intersection for each of the triangles at the different depths. Finally, a line-of-sight equation is found by fitting a straight on the different corners, as shown in figure 2.15. This calibration procedure is too slow to be applied to all the sensing elements, so in practise, a number of 25 evenly spaced of them are found. Finally, a least-squares procedure is used to fit a pinhole camera model. Experimental data for this system shows that for a range value within 0.5mm to 500mm, an accuracy of 0.1% is achieved. A transfer rate of 1000 $2\frac{1}{2}$D images per second is achieved using this system. Although the improvement in the reconstruction accuracy is very significant, it could even be improved if a real *peak detector* was used instead of a simple threshold circuit. In addition, threshold time stamp recording is strongly dependent on the object reflectivity properties, hence different time stamps are obtained for equally shaped object with different reflectivities.

---

[1]DOF:Degrees Of Freedom

2.4.2   Projective approach

Chen & Kak [CK87] developed a similar three-dimensional digitiser, but the camera-projector set was arranged so that it formed a rigid structure, keeping their relative orientations invariant to each other. The whole system was mounted in the end effector of a robot arm in order to scan the scene under consideration. Chen & Kak proposed a new calibration method based on the application of projective geometry (see figure 2.6). They stated that the points on the light plane could be mapped onto the image plane by finding the planar transformation (homography) which relates both planes, as shown in figure 2.7. The reader is pointed to [Fau93] or [HZ03] to get deeper into projective geometry. The planes **s** and **r** correspond to the laser and image planes respectively. The calibration procedure is carried out with respect to a reference coordinate frame. A different coordinate frame F_s is associated to the laser plane which describes its pose with respect to the reference frame. A bi-dimensional coordinate frame F_{2s} is associated with the laser plane, where its axis x and y coincide with the axis x and y of F_s. This is an artifact used by the authors of [CK87] in order to easily change from the bi-dimensional coordinates of points on **s** to their corresponding three-dimensional representation.

$$\rho \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} \tag{2.2}$$

Using the property of the cross-ratio invariance in projective space, one can define a 3x3 planar transformation mapping the points laying on a plane **s** into another plane **r**. Then, equation 2.2 is obtained, where ρ is a scale factor. The left hand side of this equation is the vector coordinate of a point laying on **s** (X_s), while the right hand side column vector describes the coordinates of a point laying on **r** (X_r). Considering the transformation F_s, the transformation from F_{2s} to F_s and equation 2.2, equation 2.3 is obtained (where $x_1' = u$, $x_2' = v$ and $x_3' = 1$), which maps a point on the image plane into its corresponding three dimensional point. The matrix T_{cb} in equation 2.3 is called the *conversion matrix*. From equation 2.3, the expressions for x, y and z in terms of the t_{ij} components of the conversion matrix can be obtained. Hence, a set of three linear equations is set for every image point. Since ρ is a scale factor, the 4x3 matrix T_{cb} can be simplified making $t_{43} = 1$ without loss of generality. Then four points are enough in order to solve for the 11 unknowns (t_{11} to t_{42}). A system of 12 equations with 11 unknowns is raised, but since the points must be chosen such that they are coplanar, they must satisfy equation 2.4 too, which is *de facto* implicit in equation 2.3,

meaning that the 12th equation is determined by the other 11. In practise, a set of more than 4 points is used in order to set an overdetermined system of equations obtaining a solution that best fits some minimisation criterion like a least square optimisation. Experimental results made on a simple cubic object show that the relative accuracy of the system is dependent on the distance from the object to be scanned, obtaining errors of <0.05 inch at a distance of 20 inches and <0.04 inch at 14 inches, for the measure of the width, which is the side *perpendicular* to the optical axis of the camera. In addition, when the accuracy is evaluated measuring the height, which is the side *parallel* to the optical axis of the camera, the errors are one order of magnitude bigger, obtaining <0.30 inch at 20 inch and <0.14 inch at 14 inch. Again the accurate stripe detection is of great importance for an improved measurement accuracy.

Reid [Rei96] used the same idea providing the Oxford AGV with object recognition and acquisition capabilities (see figure 2.8). In this system, instead of making the device scan the scene itself, it is mounted on a static platform and a mirror is used in order to reflect both the laser stripe and the image. Experimental data is reported showing a 1% of error relative to the measured depths. The calibration procedure was similar to [CK87], but using plane-point correspondences, instead of line-plane correspondences. Huynh [Huy97] adopted the projective model in order to calibrate a fringe projection range finder. In that work, a different 4x3 transformation matrix is obtained for every fringe. The whole set of fringes is projected at a time, so every light plane must be uniquely identified in order to set its relation to the corresponding 4x3 transform. Huynh proposed a method for obtaining point to point correspondences from known 3D points, based on the invariance of the cross-ratio. The 3D data for comparison were obtained using the SHAPE [AN87] scanner. Since a commercial scanner was used, the spatial point resolution is limited by the number of fringes this scanner is able to project. The improvement in performance could only be evaluated in terms of accuracy, since there was no possibility to measure how faster or how robust in the presence of noise it was. Jokinen [Jok99] used the projective relation between a laser plane and the image plane as a starting point for refining a laser stripe scanner calibration. Park [PDK01] used double slit scanning for minimising the occlusions appeared in conventional single slit scanning, and used two calibration matrices for reconstructing the two laser planes.
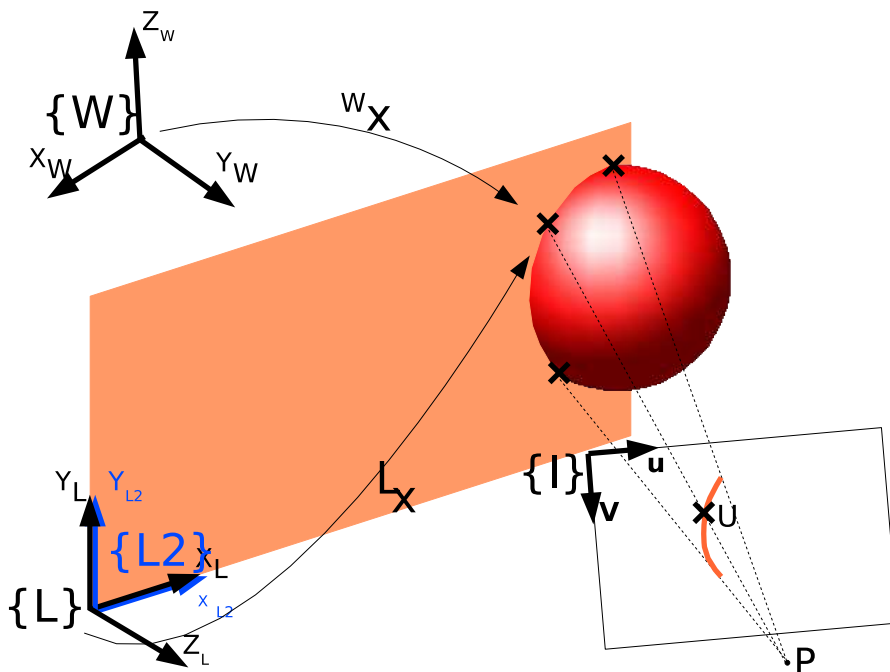
Figure 2.6: The system arrangement of Chen & Kak.

$$\rho \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = F_s \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} =$$
$$\begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \\ t_{41} & t_{42} & t_{43} \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = T_{cb} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{2.3}$$

$$\det[X_b^1 X_b^2 X_b^3 X_b^4] = 0 \tag{2.4}$$

2.4.3   Euclidean approach

Takatsuka et al. [TWVC99] developed a low cost interactive scanner. The device cast a spot on the scene by means of a LED marked laser pointer. Figure 2.9 shows the operational principle. A strong point of this system is that the scanning can be made by hand, that is, the pointer can be wiggled over
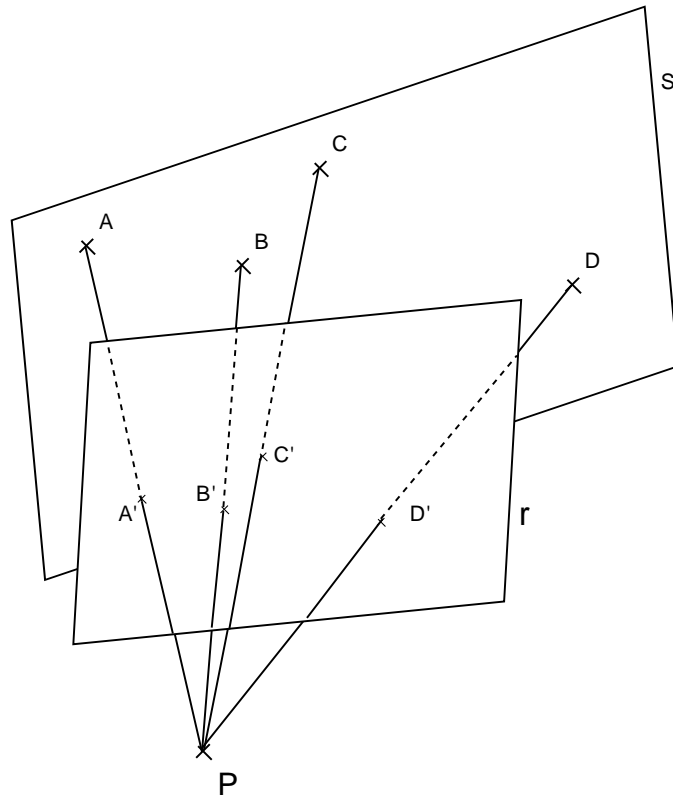
Figure 2.7: Elements of two dimensional projectivity.
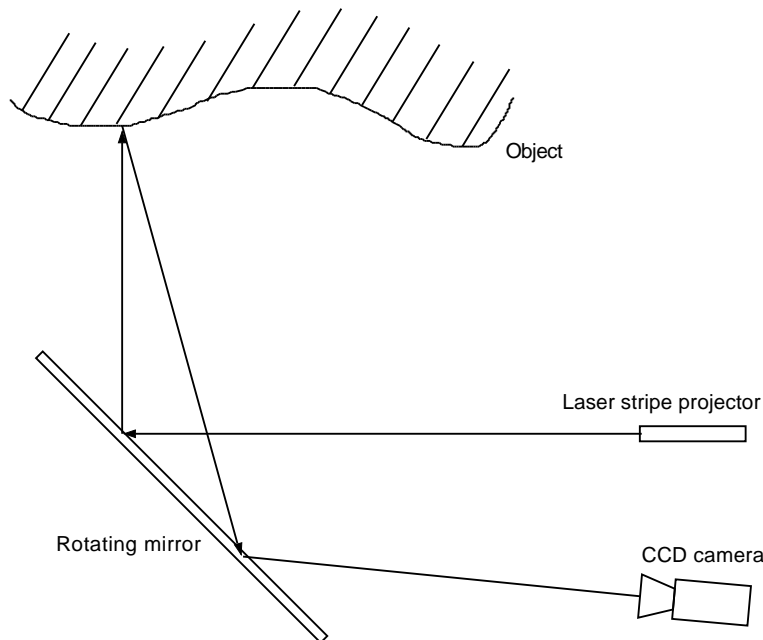
Figure 2.8: Oxford/NEL range-finder scheme.

the scene and no mechanics at all are needed. The laser pointer can be modelled by a line equation, which is identified by the three LEDs and the laser spot on the object scene. The line equation is obtained by solving a system of equations in terms of the camera projection matrix parameters, and property of the cross-ratio invariance under projective transformations. The camera has been calibrated assuming a pinhole model. The system of equations is solved by equating to zero the partial derivatives of the lagrangian, as shown in equation 2.5, where l is the line direction vector, p'_D is the 3D laser spot, and D, E are 8x3 matrices in terms of the camera projection matrix and the cross-ratio beween the four collinear points. Hence, the reconstruction of the imaged spot in 3D is straightforward equating the partial derivatives of the lagrangian with respect to l and p'_D to zero. On the contrary, since the lagrangian yields two possible solutions, there is an ambiguity in determining the solution of the 3D coordinates of the laser spot. The reconstruction accuracy is strongly dependent on the size of the spots, both the laser and the LEDs. Since the line equation is obtained by fitting a line to the three LEDs of the pointer, it is obvious that this pointer must be visible, at any time, by the camera, hence limiting the field of view. Since the three LED collineation is a critical aspect for a good precision in the determination of the line equation, it is straightforward to make sure that this collineation is
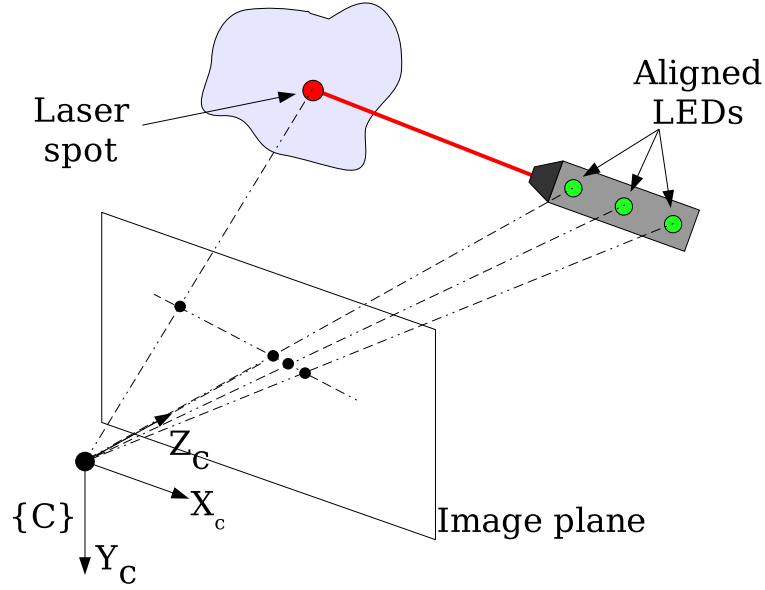
Figure 2.9: Low cost *wand* scanner.

maximised, requiring a high cost specially designed laser pointer. However, the high interactivity and the *magic* that seems to develop on the screen while reconstructing, makes the laser pointer look like a wand.

$$L(p'_D, l, \lambda) = \|Dp'^T_D + El^T\|^2 + \lambda(1 - \|l^T\|^2) \qquad (2.5)$$

Similarly, Furukawa [FK03] used a laser point projector with LED markers attached to it for use as a slit scanner. It is the natural extension of Takatsuka's proposal to a laser plane. Figure 2.10, shows a sketch of the operation principle, where four LEDs are used to define the laser plane equation. As can be seen, the laser plane is described by a three-dimensional coordinate system composed by vectors e'_1, e'_2 and e'_3, where e'_1 and e'_2 are parallel to two oposite sides of the square defined by the LEDs, and $e'_3 = e'_1 \times e'_2$. The corresponding images of the axis, $\widehat{e}'_1$ and $\widehat{e}'_2$, are estimated using the vanishing points generated by the horizontal and vertical sides of the image of the LEDs, respectively. Using these correspondences, the plane equation is estimated. The system allows a manual scanning, like in [TWVC99], but increases the reconstruction speed, since a whole slit scans the scene, instead of just a spot. The calibration consists on identifying the laser plane equation from the axis correspondences. However, a former camera calibration is needed in order to estimate the mismatch between the 4 LED-defined plane and the laser plane. Like in Takatsuka et al., the wand is required to appear
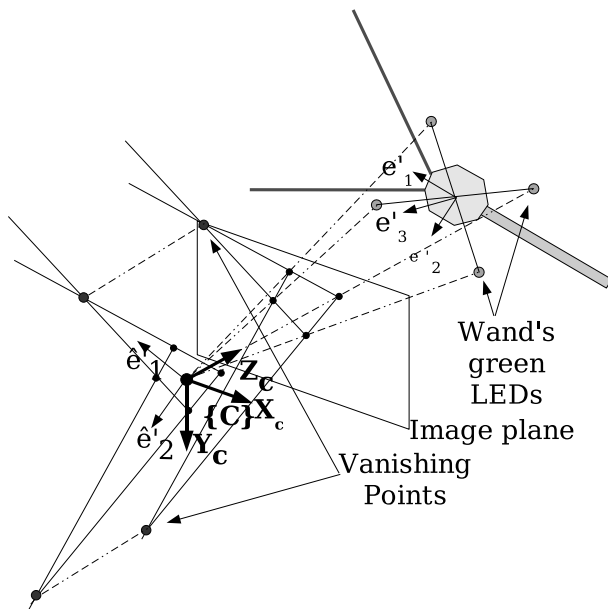
Figure 2.10: Low cost *wand* slit scanner.

on the image, since the reconstruction is impossible otherwise, and hence the field of view is limited. Since the coordinate systems identification is based on the computation of the vanishing points, instability in the calculations and hence reconstruction *gaps* can appear in case that the two sides of the square appear parallel or quasi-parallel on the image. In that case, the vanishing points tend to infinity.

Chu et al. [CHJ01] used a well known structure as a reference for calibration free reconstruction. The structure is a cube with LEDs attached to the vertices and edges, with very well known dimensions, as shown in figure 2.11. Since the LEDs locations are well known, the cross-ratio invariance under projective transformations is used in order to extract the 3D coordinates of every point on the laser stripe. The cube, which must always be visible, establishes the reference coordinate system. The cross-ratios are calculated between four points, which, similarly to Furukawa et al., one of the four points is obtained by computing the vanishing point of the lines containing the visible edges of the cube. As reasoned previously, under certain conditions, vanishing point computation can induce serious inaccuracies due to their tendency to infinity. Since the intersection between the laser plane and the cube is always visible, and hence, referred to the reference coordinate system, there is no need to compute the laser plane equation. This fact has
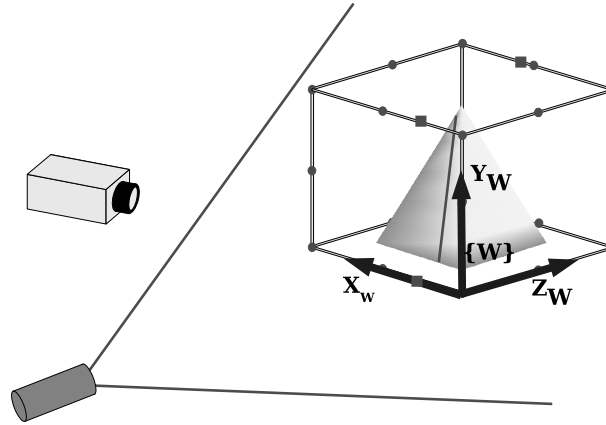
Figure 2.11: Calibration-free 3D laser scanner.

led the authors to name the scanner as a *calibration-free* approach, but the use of a well known reference structure (the cube), which has been manufactured using accurate mechanics, is indeed a form of calibration. In addition, since the cube must be visible on the image, the field of view and the size of the object is limited, and occlusions are induced in the reconstruction.

2.4.4 Time multiplexing. Switching the laser slit

Space-encoding stands for the projection of successive binary patterns onto the scene to be acquired (see figure 2.12).

The number of bits used in the codification is directly related to the number of patterns projected. Usually, the bright regions are assigned the logic level '1', and the dark regions are assigned a '0'. Finally, if each pattern is assigned a different weight, every pixel is associated to a unique binary code, which can either be natural, grey or any other suitable code. Space-encoding techniques are being applied using pattern projection ([PA82, ISM84, HK99, SCR00]). In these cases, the projector is modelled as a *reverse camera*. These and other pattern projection scanners can be found in the surveys of references [BMS98, SPB04].

Sato et al. applied this technique in the *CubicScope* three-dimensional digitiser [SO93], using a scanning laser slit mechanism. The technique takes advantage of the shutter period of a standard CCD camera: within the time period in which the shutter is open, the laser slit must scan the scene in such a way that illuminated (bright) and shadowed (dark) regions appear. This technique requires the laser to be rapidly switched on and off. Due to
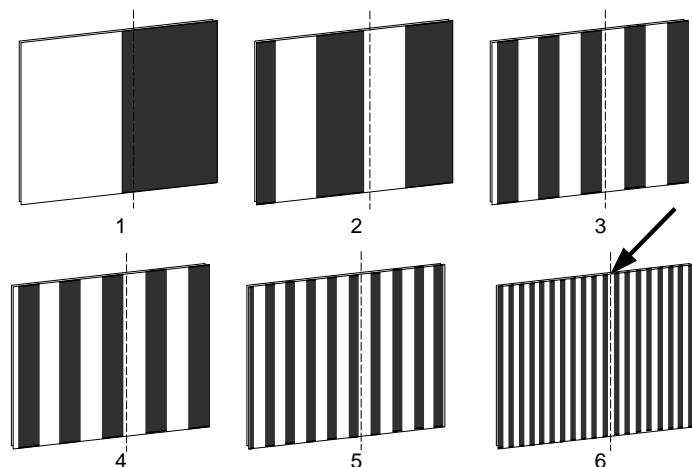
Figure 2.12: Space-encoding method.

the light integration given on the CCD, it stores a whole image as though a slide pattern was projected onto the scene. Hence, for each shutter period, a different pattern must be generated using the switching technique in order to get the proper patterned image. Generally, N encoded images produce the same range map than $2^N - 1$ non-encoded images, taking one image shot per each laser scan step.

As an example of performance, consider the operation of a standard NTSC video camera which operates at 30 frames per second. Then, in order to achieve a rate of 1 $2\frac{1}{2}$D image in 0.3 seconds, a series of 9 patterns must be projected. This implies that a $2^9 - 1 = 511$ *columns* range map is obtained with 1% of accuracy, according to [SO93]. Cubicscope is calibrated in two steps: first, the CCD camera is calibrated using Tsai's method [Tsa86]. Second, regarding the scanning nature of the laser slit projection, each of the coded regions is assigned to a plane equation, hence, the second step of the calibration process is in charge of the accurate identification of every equation. In this sense, a mechanical arrangement consisting on an accurately positioned plane at different ranges from the digitiser is set up, then, on each of the planes a whole set of patterns is projected and the images are stored. Every laser slit plane is identified by the equally coded regions imaged on each of the calibration planes, hence one laser slit plane equation is yield for every coded region. In [HS96], a new approach using pattern shifting is tested with the aim of improving the accuracy at the expense of acquisition speed. The pattern shifting method relies on the accurate detection of the regions' edges using the CCD charge storing effect.
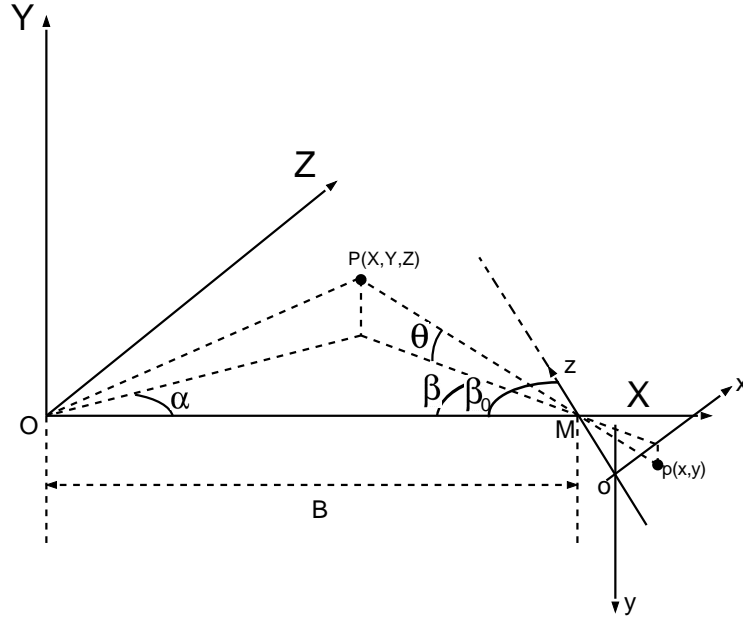
Figure 2.13: Yu et al. digitiser geometric parameters.

Yu et al. [YZWQ96] used the space-encoding method in their laser scanning device, whose mechanical arrangement is shown in figure 2.13. Expressions for X,Y and Z are obtained using trigonometrical analysis, as equations 2.6 to 2.8 show, where $2\theta_1$ and $2\beta_1$ are the vertical and horizontal view angles of the camera, respectively. The number of horizontal and vertical pixels are 2N and 2M respectively, while n and m are the corresponding pixel indexes.

$$Z = \frac{B}{\cot\alpha + \frac{\cot\beta_0 + n\cdot(\tan\beta_1/N)}{1 - n\cdot\cot\beta_0\cdot(\tan(\beta_1/N)}}$$ (2.6)

$$Y = \frac{m\cdot Z\cdot(\tan\theta_1/M)}{\sin\beta_0 - n\cdot\cos\beta_0\cdot(\tan\beta_1/N)}$$ (2.7)

$$X = Z\cdot\cot\alpha$$ (2.8)

The three-dimensional errors introduced by the error in the pixel detection (Δx and Δy) as well as the angle error ($\Delta\alpha$) are studied obtaining the expressions for ΔX, ΔY and ΔZ, by partially deriving X, Y and Z with respect to x, y and α. A study about the error introduced by the geometry of the polygonal mirror was reported (see table 2.1). In such arrangement, the rotating mirror has several different reflecting sides in order to achieve the optimal number of scans per motor turn. This makes the reflected slit to

effectively be projected from somewhere away from the rotational axis, as shown in figure 2.14, and this introduces an error in the calculation of the baseline B (ΔB_0). Equation 2.9 shows an expression for ΔB_0, deduced by [YZWQ96]. A very thorough work on the estimation of calibration errors in laser scanning devices using space-encoding can be found in [YZWL98]. This work was done from the results published in [YZWQ96]. In [YZWL98], Yu's digitiser is calibrated using a different method. The mechanical arrangement consists of a standardised plane, placed in front of the lenses, within the view of the system, such that it is parallel to the *xoy* plane and perpendicular to the *xoz*. Using the switching technique for pattern generation, two concentric, different size squares are projected onto the calibration plane. Then, the plane is moved along the z axis until the biggest square fills the whole image plane, the calibration plane position is recorded, and the process is repeated for the smallest square. Using the known geometrical relation between the two squares, the different parameters are obtained. The reconstruction is done using triangulation and an accuracy of 1.5% is achieved.

In space-encoding, the *width* of each coded region determines both the resolution and the reconstruction accuracy. Hence, the number of bits involved in the pattern generation must be carefully chosen according to the application specs. In addition, a unique criterion must be systematically applied in order to determine which of the several subsets of points laying on a coded region is assigned to each plane equation with the aim of minimising the stripe detection uncertainty as well as the reconstruction inaccuracies.

$$\Delta B_0 = 2R \cdot (\cos a - \cos a_0) \cdot \frac{\sin(b - a)}{\cos(b - 2a)} \qquad (2.9)$$

2.4.5   Lookup Table generation

Yokohama et al. [YSYI94, SYI94b] proposed a new prototype similar to that of [GKC90], incorporating notable improvements in slit detection. They used two side by side photosensitive areas in every cell, sensing the *difference* in light intensity between the two twin photo-diodes in every cell. The cell architecture uses three clock signals in order to set a 4-phase operation for synchronously sensing and transferring the data. A notable improvement in the accuracy has been achieved, obtaining a range error of -24μm to +14μm, and is, in addition, very robust under the presence of significant levels of ambient light. Hori et al. [HBK94] and Baba et al. [BKK97] proposed similar systems based on the *cell-parallel* processing concept, introduced by [ASP87]. Recently, Brajovic and Kanade [BMJ01] have proposed a new approach to
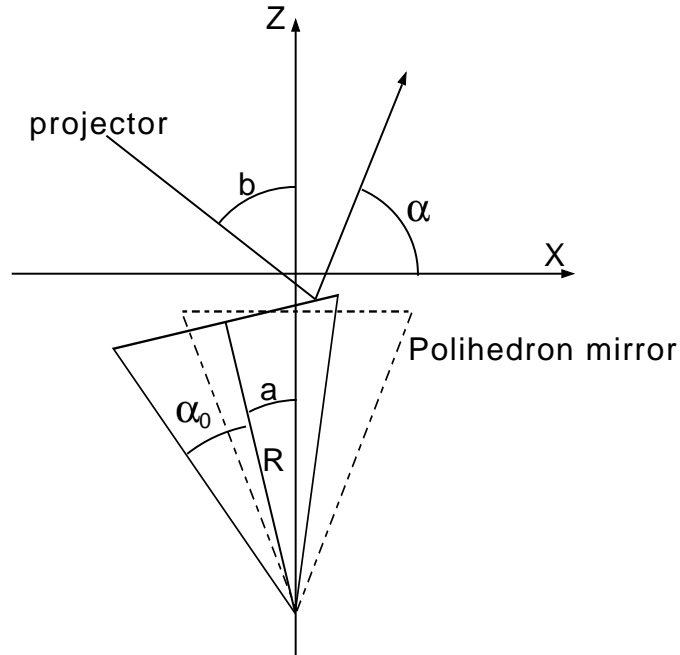
Figure 2.14: Change on the slit ray projection axis.

VLSI acquisition of $2\frac{1}{2}$D images, introducing the concept of *row-parallel* processing. This new idea takes advantage of the fact that laser slit scanning is inherently row-parallel, since only one cell on each row is illuminated at a particular instant of time.

Yokohama et al. used a plane object for calibration. The procedure consists of 4 scans over the same plane, using 4 different depths. In each scan, the relationship between the scanning angle and the plane depth is recorded for each photo-sensitive cell. At the end of the process, a lookup table is generated in order to allow for the real-time three-dimensional reconstruction. The fact that stripe detection is made by comparison between two side-by-side photosensitive areas in each cell yields a significant improvement in the accuracy. Several considerations must be taken, since the signal-to-noise ratio of each electric signal must be maximised in order to get the highest reconstruction accuracy. To this aim, a tradeoff must be achieved among scanning speed, photosensitive area and optical laser power in order to successfully comply with the application specs. In addition, since time is the variable used for measuring the scanning direction, it is clear that high resolution timers and high clock frequencies must be used, hence the
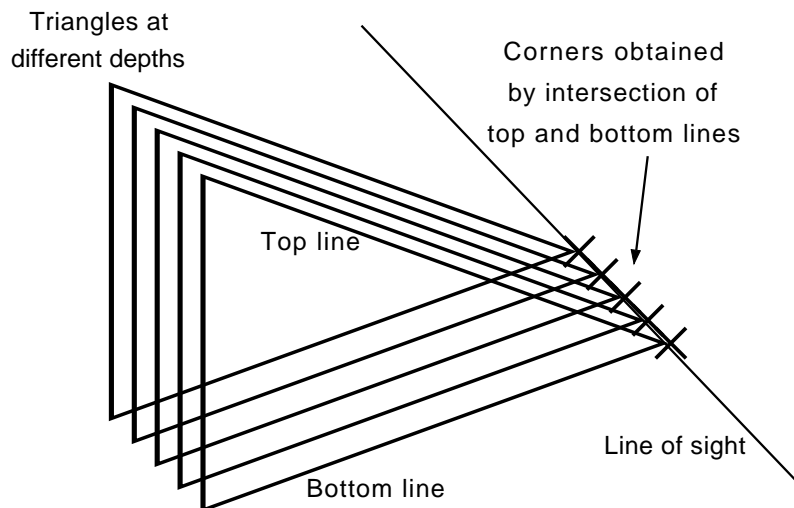
Figure 2.15: Line of sight computation in Gruss et al.

upper limit in scan resolution is the highest clock frequency that can be used.

Even when this approach is straightforward and easy to implement, few authors have reported results, and hence no other improvements can be explained.
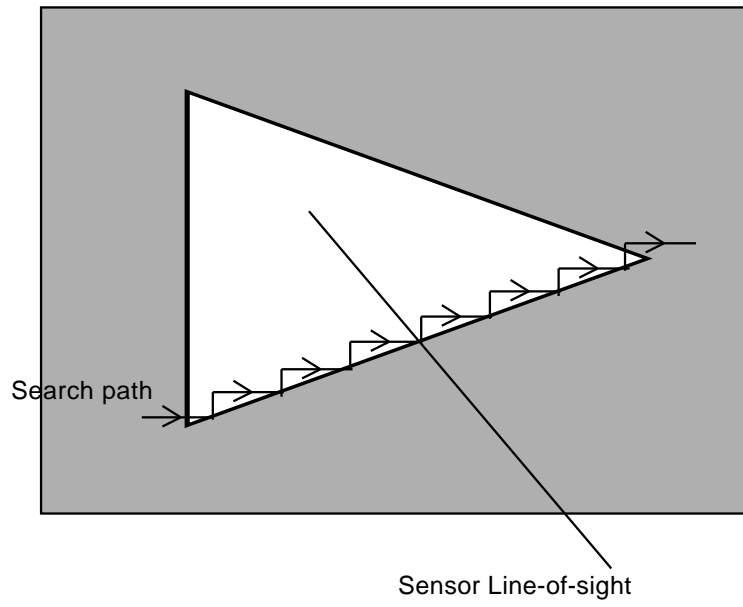
Figure 2.16: Calibration target used in Gruss et al.

Chapter 3

Laser peak detection

The accuracy of a 3D reconstruction using laser scanners is significantly determined by the detection of the laser stripe in the image. Since the energy pattern of such a stripe corresponds to a Gaussian profile, it makes sense to detect the point of maximum light intensity (or peak) by computing the zero-crossing point of the first derivative of such Gaussian profile. However, because noise is present in every physical process, such as electronic image formation, it is not sensitive to perform the derivative of the image of the stripe in almost any situation, unless a previous filtering stage is done. Considering that stripe scanning is an inherently row-parallel process, every row of a given image must be processed independently in order to compute its corresponding peak position in the row. This chapter reports on the use of digital filtering techniques in order to cope with the detection of a laser stripe for the scanning of different surfaces with different optical properties and different noise levels, leading to the proposal of a more accurate numerical peak detector, even at very low signal-to-noise ratios. However, the results of the experiments carried out in our lab are reported in chapter 5, and they are accompanied by a comparison with other five existing peak detection methods.

3.1   Introduction

The reconstruction accuracy depends on a set of cross-related issues like calibration [SAB02], camera resolution, optics distortion, noise [Cur97, Die00], etc, while the range acquisition time is dependent on a smart image processing algorithm (responsible of segmenting the appropriate regions of interest) in addition to a fast imaging sensor. In this chapter, a numerical peak detector based on the computation of the zero-crossing of the first derivative of

each image row is analysed. Similarly to the estimator proposed by Blais and Rioux [BR86], a derivative filter is used for computing the first derivative, but its coefficients and order are selected according to the surface optical properties and a previous frequency analysis [HSG00, Sar91, Suc91] of the image. The performance in terms of peak location error of this method is similar to that of [BR86] when the signal-to-noise ratio of the stripe image is high, but it is significantly improved when the nature of the surface or a low light power induce a high noise level in the stripe image [GRB+01]. This chapter is structured as follows: the next section explains the different surface properties under consideration. The noise sources which affect the 3D measurement are briefly explained in section 3.3. Section 3.4 analyses the proposed method.

3.2   Scanning different surfaces

The optical properties of the surface significantly determine the performance of the laser scanner. The optimal surface type for scanning purposes is a totally lambertian surface with a high reflection index. Figures 3.1 show how a light ray behaves under both a specular and a lambertian surface. Translucid surfaces are often present in our everyday life (certain types of plastic, animal tissue, silicon, resins, certain rocks or minerals, etc.). Figure 3.2 shows how a ray of light behaves when it impinges such kind of surface. In a translucid surface, light reflects as in a lambertian surface, but it goes through the material until a certain depth. The higher the light power, the deeper the light penetrates inside the material. In addition, the light scatters inside the material, so that a camera looking at it "sees" laser reflexions sourcing from inside it. See [GRB+01] for an example of light behaviour and a study of how it affects the 3D measurement on marble surfaces. Figure 3.3 (right), shows a laser stripe reflected on a lambertian surface, while figure 3.3 (left), shows how the reflection on a translucid surface is seen by the camera. As it is shown, a laser stripe impinging on a translucid surface induces a lot of undesired light peaks where they are not expected to be. In addition, if the light power is lowered, the noise due to the different sources becomes more and more significant and hence, the reconstruction quality degrades.

3.3   Noise sources

In standard cameras, the most influencing noise has been found to follow a Gaussian probability distribution, which is a consequence of the point spread
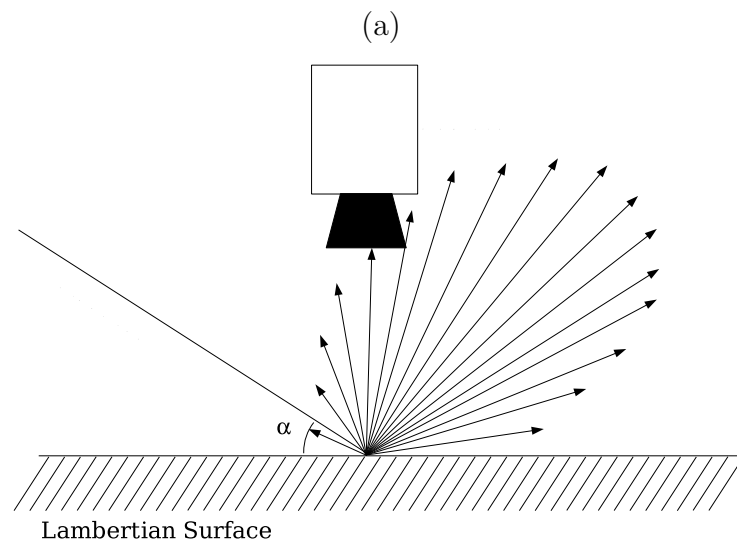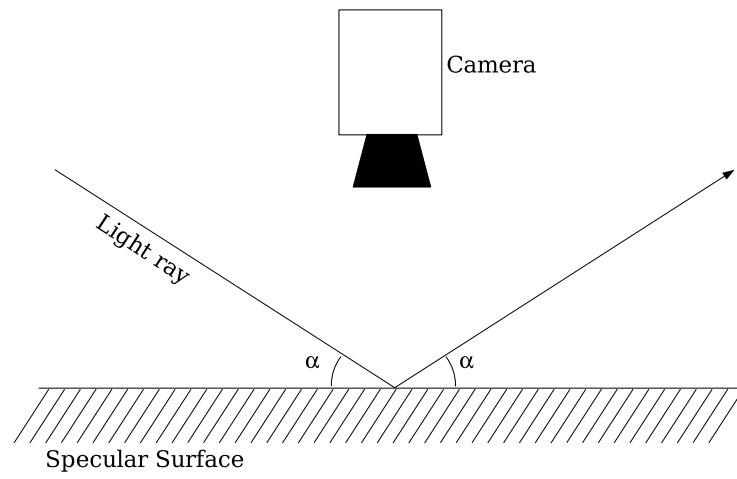
Figure 3.1: Behaviour of light reflected on a specular surface (a), and a lambertian surface (b).
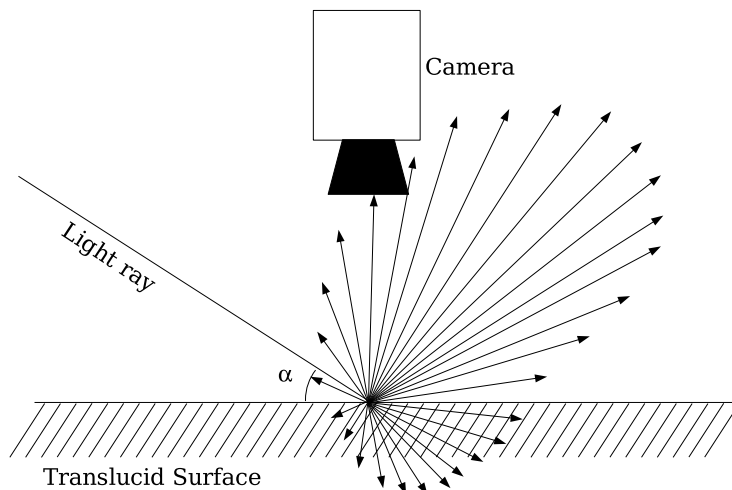
Figure 3.2: Behaviour of light reflected on a lambertian surface.

function, due to the imperfections in the lenses and the grey level digitisation. Three noise sources have been found to influence the three-dimensional measurement of camera-laser based 3D scanners: electrical noise, quantisation noise and speckle. The speckle is directly related to the nature of laser light, while the other two noise sources are inherent to the image sensor. Electrical and quantisation noise are very significant when the S/N[1] is very low, i.e. when the stripe light power is very low. A thorough study about the performance of CMOS image sensors and how electrical noise affects the image quality can be found in [Die00]. This is commonly found when the surface exhibits a very low reflection index or when scanning at high speeds. In addition, speckle noise influences dramatically on the measurement. Speckle is due to the reduced wavelength of light compared to the surface roughness and the monochromacity of laser light [Cur97] and influences the 3D measurement. These three noise sources, combine together and make the observer see the constructive and destructive interferences within the laser stripe.

---

[1]S/N: Signal-to-noise ratio
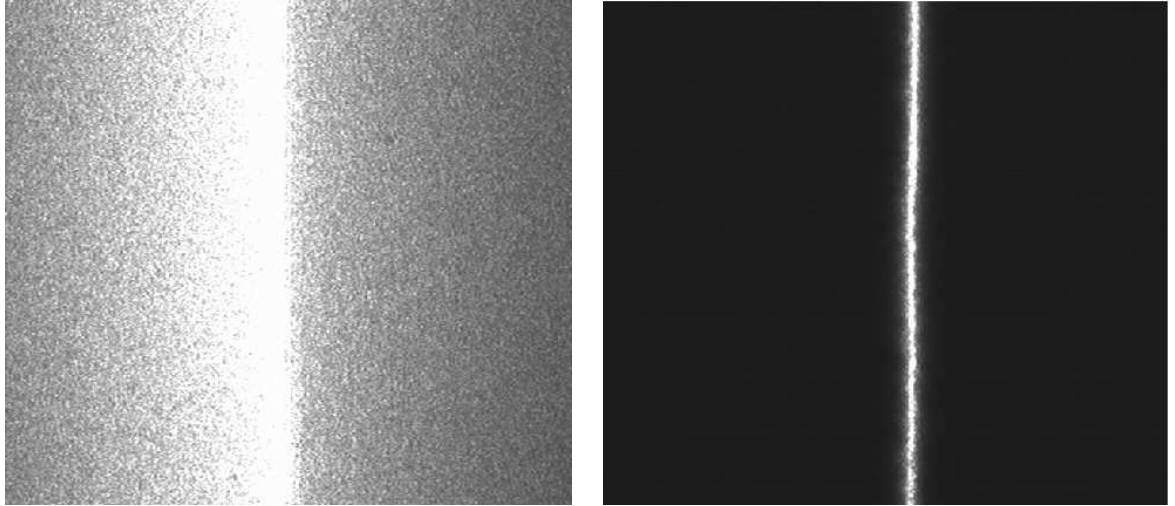
Figure 3.3: A laser stripe on a translucid (left) and a lambertian (right) surface.

3.4 A new method for peak detection

Either due to any of the noise sources explained in the previous section or the scattering of light inside a translucid material, the camera images a laser stripe with a certain amount of undesired lighting peaks superimposed to it. From the point of view of signal processing, it seems reasonable to consider it as the manifestation of a noisy signal, which complies with the principle of superposition. Regarding each row of the stripe image as a signal, a digital low pass filter can be designed with the right cut-off frequency, attenuation and transition band width parameters. There are several methods for obtaining the filter coefficients. Haddad et.al. [HSG00] developed a new technique for designing FIR[2] filters based on the method of vector-space projection. Other approaches include the method of adjustable windows [Sar91] or the use of genetic algorithms for certain types of FIR filters [Suc91]. Overall, the estimators presented in the literature are very sensitive to the variations in S/N, seriously constraining their performance on many types of materials, especially considering translucid surfaces. In this work, the computation of the zero crossing point of the first derivative of each stripe image row has been considered for obtaining an estimation of the peak position. Although other methods use the same approach, in the present work much more emphasis has been put in the obtention of the derivative operator, generalising

---

[2]FIR: Finite Impulse Response

the notation in terms of the *convolution* operation. As stated in equation 3.1, FIR filtering consists on computing the convolution of the row signals $(x(n))$ with the coefficients of the filter $h(n)$. As it has been shown, the first derivative of the convolution of $x(n)$ and the filter coefficients $h(n)$ is equal to the convolution of $x(n)$ and the first derivative of $h(n)$. Once the filtered first derivative has been computed, assuming that the S/N is high enough for differentiating the laser stripe from the rest of the scene by simple thresholding, the method for obtaining the zero crossing point is as follows. Formerly, the maximum grey level value is selected. Second, the signal is tracked from its maximum, left to right in the image, until the first negative value is found. Finally a straight line is computed between the points corresponding to the first negative and the last positive signal values, and the zero crossing is computed. Figure 3.4 summarises this process and equation 3.3 shows how the estimation of the zero crossing is computed.

$$y(n) = \sum_{k=0}^{M} x(k) \cdot h(n-k) \triangleq x(n) * h(n) \tag{3.1}$$

The first derivative of the filtered row, $\frac{dy}{dn}$, can be computed as shown in equation 3.2, according to the properties of convolution.

$$
\begin{aligned}
\frac{dy(n)}{dn} &= \frac{d[x(n)*h(n)]}{dn} = \frac{d\sum_{k=0}^{M} x(k) \cdot h(n-k)}{dn} = \\
&= \sum_{k=0}^{M} \frac{d[x(k) \cdot h(n-k)]}{dn} = \sum_{k=0}^{M} \frac{dx(k)}{dn} \cdot h(n-k) + \\
&+ \sum_{k=0}^{M} x(k) \cdot \frac{dh(n-k)}{dn} = \sum_{k=0}^{M} x(k) \cdot \frac{dh(n-k)}{dn} = \\
&= x(n) * \frac{dh(n)}{dn}
\end{aligned}
\tag{3.2}
$$

$$\hat{X} = x_0 - \frac{y_0 \cdot (x_1 - x_0)}{y_1 - y_0} \tag{3.3}$$

3.5   Summary

A new method for laser stripe peak detection based on the computation of the first derivative zero crossing has been explained. The computation of the first derivative zero crossing of a function is a very well known technique for obtaining local maxima and minima, and it is taught in basic mathematics courses. The problem of applying this technique to the field of laser stripe scanning is that noise is a big issue to deal with. As it has been shown,
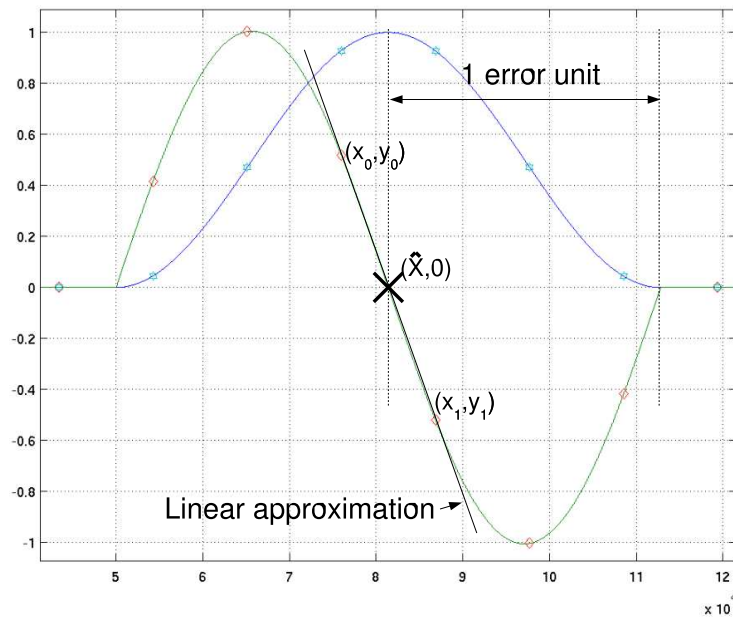
Figure 3.4: Laser peak and first derivative.

several noise sources influence in the peak detection, and different surfaces affect the results significantly. Obviously, if the noise could be removed completely, the problem would have gone, but the best one can do with noise is to try to minimise as much as possible through filtering techniques. In this proposal, FIR filters have been used, and an important property of convolution has been used in order to filter and derive in a single convolution. The backbone for a successful filtering is to obtain the right filter coefficients and derive them, and this step has been done empirically by observing the FFTs of the stripe image rows, selecting the right cut-off frequency and the right transition band width. It is clear that the automation of this process would be more than welcome, and is a clear step which must be done as a further work. Nevertheless, the results that where obtained in laboratory experiments have been very encouraging, and they can be checked in chapter 5.

Chapter 4

Calibration with the Complete Quadrangle

Three-dimensional scanners are used for measurement purposes. They measure shapes by acquiring (or digitising) a set of points, commonly referred to as a *cloud of points*, which is a discrete approximation of a real object. Obviously, the dimensions of the digital representation should be as close as possible to the dimensions of the real object. One of the very well known quantities used in any metric measurement process is *accuracy*, which is defined as the maximum radius of the sphere inside which there is total probability of finding the coordinates of a measured point. From the above definition, it is clear that the smaller this radius is, the better the approximation of the reconstruction to the real object will be. The key factors for maximising accuracy are essentially the method for obtaining the laser light peak and the calibration method. This chapter describes an elegant projective calibration method that obtains the best 2D to 3D point correspondences taking advantage of the invariance of the cross-ratio under collineations and a specially designed calibration target based on the shape of a complete quadrangle.

4.1 Introduction

The projective calibration of cameras rely on the identification of a good set of 2D to 3D point correspondences, which are used for obtaining **a)**the camera matrix P for single view geometry [HZ03, SAB02] , **b)**the fundamental matrix F for two view geometry [OF86, AS03] and **c)**the trifocal tensor T for three-view geometry [Fau93, FP97] . Tensor notation can be expanded to deal with the problem of N-view geometry based on the identification of point correspondences, too [HZ03]. Each of the above problems is reduced to a pa-

rameter estimation problem, which can be solved with state-of-the-art min-imisation techniques [CBvdHG00, CBvdHG01, CBvdHG03, KOK00, LM00] either directly or iteratively. It is well known that the better the point cor-respondences match, the better the parameter estimation, and hence the system can be modelled more reliably. One of the contributions of this work is a method for obtaining the best 3D correspondence to previously cho-sen 2D points on the image. Since the 3D points used for calibration lie on the laser stripe, an explicit calibration of the camera is avoided and the whole camera-laser emitter system is modelled by a 4x3 matrix defined up to scale with 11 degrees of freedom, as stated by Chen [CK87]. However, the technique can also be used for camera calibration. Moreover, a contribution toward the extension of Chen's method to systems in which the laser stripe scans the scene, while the camera does not vary its orientation relative to the object being scanned, is provided. A 4x3 matrix is obtained for each scanning calibration position and a function interpolation is used for every matrix parameter, in order to obtain a high scanning resolution. The method has been simulated for both angular and linear scanning, and an experimen-tal linear scanning system has been implemented. This chapter is structured as follows: the next section briefly describes the previous work, with regard to the projective calibration of laser scanners. The proposed method for ob-taining point correspondences is described in section 4.5, while section 4.6 shows how the projective method is used for scanning lasers. In section 5.3 we give the experimental results.

4.2 Projective geometry

The term *projective*, or *perspective* was introduced by the painters of the Italian Renaissance. Great names like Leonardo, Raphael or Michelangelo appear linked to such concepts. Their pictures and drafts looked so realistic that they seemed to fly outside their frame, because they got the right optical effect that made a planar picture portray a scene as if it was an open window. The optical effect was related to the observation of reality into the distance, that is, parallel lines seem to intersect at the infinity, at the point called *vanishing point*. Objects appear smaller, as they are more distant from the observer. Figure 4.1 shows The Last Supper, by Leonardo da Vinci. It is a sample of such geniality, which has been formalised mathematically in a number of forms. In the picture, the left and right edges of the roof seem to converge, which is just the impression that an observer would have if the picture was a real scene. Cameras image the world in this way, that is, the picture by Leonardo is exactly like a picture taken by a projective camera,

and hence, still today we use to qualify the Renaissance pictures as realistic.

4.2.1 Homogeneous coordinates

In a projective pinhole camera, all rays of light impinging into it, pass through the focal point. Hence, all points laying on the same ray of light cast on the same point on the image plane. The direction of these rays are important, the distance to the point that generates them is irrelevant in the image formation. In this sense, one can think that a point on the image plane corresponds to a line in 3-space, which contains both the focal point and the image point. In projective geometry, all points contained in the same line are equivalent, and this is more easily expressed using homogeneous coordinates, as stated in equation 4.1. A point in the projective plane ($\mathbb{P}^2$) is represented with a 3-vector of $\mathbb{R}^3$. In addition, any point in homogeneous coordinates has its corresponding euclidean representation, with the third component equal to 1, as shown in equation 4.2.

$$[X, Y, T]^T \sim [\lambda X, \lambda Y, \lambda T]^T \tag{4.1}$$

$$[x, y, 1]^T = [X/T, Y/T, 1]^T \tag{4.2}$$

Note that a point in $\mathbb{P}^2$ with the third component equal to 0, expresses points at infinity. There is no way to distinguish between two points at infinity using euclidean coordinates. On the contrary, homogeneous coordinates are useful to this end, and this feature is used for modelling vanishing points in computer vision.

There is an interesting property in projective space, which is the so called principle of duality. This principle states that a geometric figure has its dual in another geometric figure. For example, lines in $\mathbb{P}^2$ are the duals of points, and in $\mathbb{P}^3$, the duality exists between planes and points. Not surprisingly, points and lines have the same number of components in $\mathbb{P}^2$, and the same happens with points and planes in $\mathbb{P}^3$. Equation 4.3 shows the relationship between a line $l = [a, b, c]^T$ and a point $p = [X, Y, T]^T$, that is, a point p in $\mathbb{P}^2$ belongs to a line l if and only if their dot product vanishes.

$$l^T p = aX + bY + cT = 0 \tag{4.3}$$

The line l, expressed in homogeneous coordinates, that contains both points p_1 and p_2 can be found by computing the cross product of p_1 and p_2 (4.4), and by duality, the point p, which is the intersection of lines l_1 and l_2, can be obtained by computing the cross product of l_1 and l_2.

$$\begin{aligned} l = p_1 \times p_2 \\ p = l_1 \times l_2 \end{aligned} \tag{4.4}$$
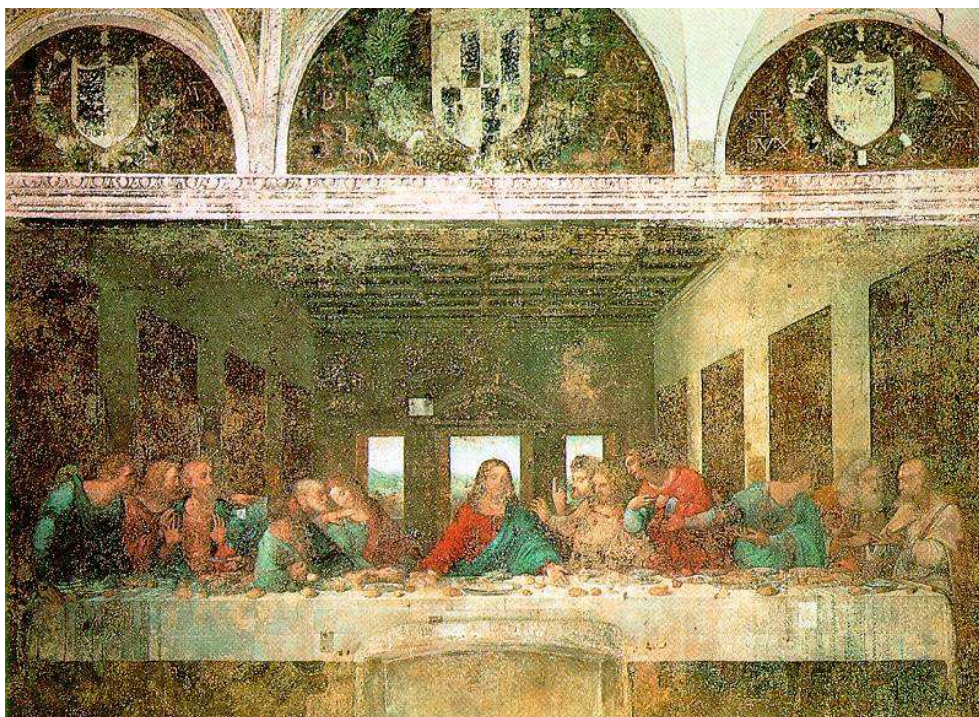
Figure 4.1: The Last Supper, by Leonardo da Vinci

4.2.2 The cross-ratio

The cross-ratio of a pencil of four lines can be defined by the four points of intersection of a fifth line not pertaining to the pencil. If these four collinear points are labelled A, B, C and D, the cross-ratio can be defined by equation 4.5. It can be proven that under projective transformations of points A, B, C and D, the cross-ratio is invariant. Figure 4.2 shows such a situation. Since the cross-ratio is a ratio of ratios of lengths, it is straightforward to express the line equations in their parametric forms. In addition, if the line equations are computed using the first and the fourth points, i.e. points A and D , according to equation 4.6, $\lambda = 0$ yields $r = A$, $\lambda = \overline{AB}$ yields $r = B$ and so forth. Hence, using the parametric forms of the line equations in this way, $\lambda_{1..3}$ can be defined to take the values: $\lambda_1 = \overline{AB}$, $\lambda_2 = \overline{AC}$ and $\lambda_3 = \overline{AD}$. According to these definitions, the cross-ratio of points A,B,C and D can be computed as shown in equation 4.7. Under projective transformations, collinearity is preserved, and the value of the cross-ratio remains invariant. Hence, the cross-ratio of points A to D is the same as the cross-ratio of points A' to D' in figure 4.2. In addition, if the cross-ratio and only 3 of the 4 points are known, equation 4.8 can be used for obtaining the coordinates of the fourth point. The cross-ratio of a pencil of four planes can be computed in $\mathbb{P}^3$.

$$Cr\{A, B; C, D\} = \frac{\overline{AC} \cdot \overline{BD}}{\overline{AD} \cdot \overline{BC}} \tag{4.5}$$

where $\overline{XY}$ is the distance between points X and Y

$$r = \lambda \cdot \frac{D - A}{\|D - A\|} + A \tag{4.6}$$

$$k = \frac{\lambda_2 \cdot (\lambda_3 - \lambda_1)}{\lambda_3 \cdot (\lambda_2 - \lambda_1)} \tag{4.7}$$

$$\lambda_2' = \frac{k \cdot \lambda_3' \cdot \lambda_1'}{(k - 1) \cdot \lambda_3' + \lambda_1'} \tag{4.8}$$

4.2.3 The complete quadrangle

Any set of four non-aligned points A, B, C, D on the plane can be joined pairwise by six distinct lines as shown in figure 4.3. This four-edged figure is called the *Complete Quadrangle* and exhibits several useful properties in computer vision applications. As can be seen, this figure defines a set of 7 points: A,B,C,D, the corners of the figure, and E, F and G, which can be obtained by computing the intersections of lines AC and DB, AD and BC, and AB and DC, respectively. The complete quadrangle exhibits the following
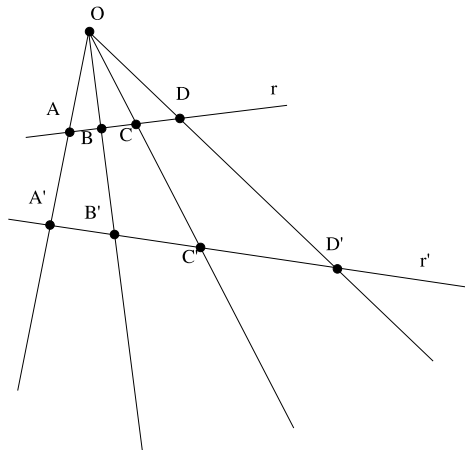
Figure 4.2: Cross-ratio of a pencil of lines

property: *The pencil of four lines based at each intersection of opposite sides is harmonic*, that is, for example, $Cr\{FA, FB; FE, FG\} = -1$. Finally, it can be seen that a rectangle can be modelled as a complete quadrangle with points F and G being points at infinity.

The essential issues on projective geometry which have been used in the development of an important contribution of this thesis, have been reported so far. There exists a wide variety of projective geometry literature, either in books, articles, unpublished or web documents. A good source of inspiration to get "involved" in the use of projective geometry for computer vision is the tutorial given by Möhr and Triggs [MT96].

4.3   Parameter estimation

A number of computer vision algorithms involve the computation of a series of parameters in the presence of noise, or noisy measurements. These parameters can be the coefficients of an equation or the components of a matrix. For example, the projection matrix of a pinhole camera, which maps the points in 3D world into their corresponding points on the image plane [Fau93]. The computation of the fundamental or the essential matrices [AS03, CBvdHG01] for solving the correspondence problem in a stereo rig, or the computation of homographies [HZ03, KOK00]. Another important field where parameter estimation is required is in curve and surface fitting [CBvdHG00, Fit97, WFAR99], which is widely used in the mathematical de-
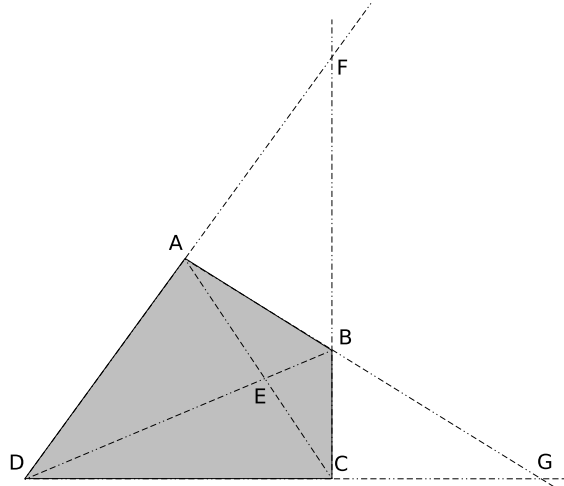
Figure 4.3: The complete quadrangle

scription of 3D surfaces. In essence, all of the above problems can be regarded as the obtention of an optimal solution to an overdetermined system of equations. This can be formalised as shown in equation 4.9, where, obviously, the trivial solution $\theta = \mathbf{0}$ is generally avoided.

$$A\theta = 0 \tag{4.9}$$

The problem stated above can be solved by direct, iterative or robust techniques. In this section, a deep comparison between different such techniques is avoided, but the reader is pointed to [HZ03], where a good overview and discussion about their performance and the field of application is addressed. Besides, only the description of a direct and an iterative method is undertaken, since these methods are used for the estimation of the calibration matrix parameters. Robust estimators are well known algorithms used for estimating parameters in the presence of outliers, or false matchings. Since the estimation of the calibration parameters is not subject to the influence of such ouliers, the robust estimator methods are not considered.

Total Least Squares (TLS)

TLS is a direct method, which consists in obtaining the linear function that best fits to a number of points, minimising the orthogonal distances of such points to the linear function, besides the Algebraic Least Squares (ALS)
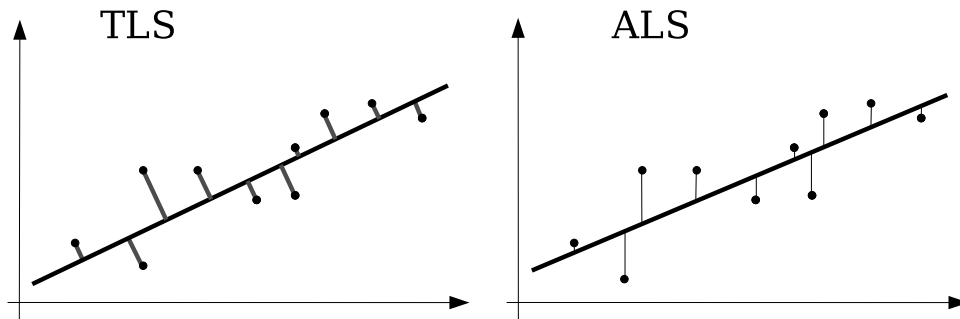
Figure 4.4: Total Least Squares and Algebraic Least Squares examples

which fits the line by minimising the algebraic distances to the points. Figure 4.4 shows this concept for a bi-dimensional linear fitting. The line equations obtained in both cases is not generally equal. TLS has proven to be much more stable and accurate than ALS [AS03], and it presents a higher robustness to the loss of rank of matrix A. One way to get a solution of equation 4.9, other than the trivial one, is by including the constrain $\|\theta\| = 1$. Hence it is the unity norm solution, which is obtained by computing the eigenvector corresponding to the eigenvalue closest to zero, in absolute value, of the square matrix $A^T A$. In case that the system of equations is not overdetermined, that is, there is the same number of equations and unknowns, the unit norm solution is found in a similar way, but using the matrix A directly, since it is already square.

Fundamental Numerical Scheme (FNS)

Iterative methods for solving sets of nonlinear equations include the yet classical Newton-Raphson or Levenberg-Marquardt methods, the method of Kanatani or renormalisation method [Kan96], the HEIV (Heteroscedastic Errors In Variables) method, by Leedan and Meer [LM00], and the Fundamental Numerical Scheme, which has proven to perform as good as Levenberg-Marquardt, is easy to implement, faster in convergence and more robust than most of the methods available.

The Fundamental Numerical Scheme was introduced by Van den Hengel et. al. [dHBCG00], as an approximation to a Maximum Likelihood Estimation problem. The method has been extensively used and compared to other methods [CBvdHG03], such as the Levenberg-Marquardt, reporting a high performance, stability and speed of convergence to the optimal solution. In addition, the performance of FNS has proven to be as high as that

of Levenberg-Marquardt, showing a significantly easier implementation.

The matrix equation 4.9, which is the basis of any homogeneous system of equations, can be rewritten as shown in equation 4.10 so that the dependence of A with the data is stated. This form of the equation is equivalent to the former one with $u(x) = A^T$.

$$\theta^T u(x) = 0 \tag{4.10}$$

Here, $\theta = [\theta_1, \theta_2 \ldots, \theta_l]^T$ represents the vectors of parameters to be estimated, $x = [x_1, x_2, \ldots, x_k]^T$ is a vector element of the data, that is, if 2D⇔3D point correspondences are used with $[u, v]^T$ and $[X, Y, Z]^T$ being their corresponding coordinates, each vector component can be $x = [X, Y, Z, u, v]^T$, and k is equal to 5. Finally, $u(x)$ is defined as $u(x) = [u_1(x), u_2(x), \ldots, u_l(x)]^T$, such that each component of $u(x)$ is a quadratic form in $[x^T, 1]^T$.

The solution θ is obtained from data observations. Since all the data obtained by imaging processes is noisy, $\theta^T u(x) = 0$ must be an overdetermined system of equations in order to obtain the best possible solution. Van den Hengel proved that the cost function to be minimised, obtained from the Approximated Maximum Likelihood Estimator (AML), in terms of the definitions above is as shown in equation 4.11. The best estimation of θ using AML is referred to as $\hat{\theta}_{AML}$.

$$J_{AML}(\theta; x_1, \ldots, x_n) = \sum_{i=1}^n \frac{\theta^T u(x_i) u(x_i)^T \theta}{\theta^T \delta_x u(x_i) \delta_x u(x_i)^T \theta}$$

where **n** is the number of data vector elements, and

$$\tag{4.11}$$

$$\delta_x u(y) = \begin{bmatrix} \frac{\delta u_1}{\delta x_1}(y) & \cdots & \frac{\delta u_1}{\delta x_k}(y) \\ \cdots & \cdots & \cdots \\ \frac{\delta u_l}{\delta x_1}(y) & \cdots & \frac{\delta u_l}{\delta x_k}(y) \end{bmatrix}$$

Letting $A_i = u(x_i)u(x_i)^T$ and $B_i = \delta_x u(x_i)\delta_x u(x_i)^T$, equation 4.11 can be rewritten as shown in equation 4.12, and since $\hat{\theta}_{AML}$ is a minimiser of 4.12, it satisfies equation 4.13, being $\delta_\theta J_{AML}$ the row vector of the partial derivatives of J_{AML} with respect to θ. This term is called the *variational equation* in [dHBCG00].

$$J_{AML}(\theta; x_1, \ldots, x_n) = \sum_{i=1}^n \frac{\theta^T A_i \theta}{\theta^T B_i \theta} \tag{4.12}$$

$$\delta_\theta J_{AML}(\theta; x_1, \ldots, x_n) = 0^T \tag{4.13}$$

In can be proved that $[\delta_\theta J_{AML}(\theta; x_1, \ldots, x_n)]^T = 2X_\theta \theta$, being X_θ a symmetric matrix, as defined in equation 4.14.

$$X_\theta = \sum_{i=1}^n \frac{A_i}{\theta^T B_i \theta} - \sum_{i=1}^n \frac{\theta^T A_i \theta}{(\theta^T B_i \theta)^2} B_i \tag{4.14}$$

And hence, equation 4.13 can be written as shown in equation 4.15.

$$X_\theta \theta = 0 \tag{4.15}$$

If covariance information is available for each data sample, the estimation improves significantly, as stated by [BCGdH00, BCGvdH01]. In this case, the expression for B should be modified as shown in equation 4.16.

$$B_i = \delta_x u(x_i) \Lambda_{x_i} \delta_x u(x_i)^T \tag{4.16}$$

Either with or without covariance matrices, FNS has proven to perform like the Levenberg-Marquardt algorithm extensively. By nature, FNS must be applied iteratively, until a convergence condition is achieved. Usually, one's desire is to achieve the the $\hat{\theta}_{AML}$ that makes $\|\hat{\theta}_i - \hat{\theta}_{i-1}\| < \epsilon$, with ϵ being an arbitrary value, as small as possible. The numerical algorithm of FNS is as follows:

1. Set θ_0 to a good initial estimate

2. Assuming that θ_{k-1} is known, compute the matrix $X_{\theta k-1}$

3. Compute the normalised eigenvector of $X_{\theta k-1}$ corresponding to the eigenvalue closest to zero in absolute value, and take this eigenvector for θ_k

4. If $\|\hat{\theta}_i - \hat{\theta}_{i-1}\| < \epsilon$ for the ϵ of your choice, then terminate the procedure. Your solution is $\hat{\theta}_{AML} = \theta_i$; otherwise, increment k and return to step 2

4.4   Previous work

The geometric relationship between two planes in projective space can be modelled by a 3x3 transformation matrix called homography. Chen and Kak [CK87] borrowed this result and used a homography for modelling the geometric relationship between the laser plane and the image plane. However, a homography relates 2D points on a plane to 2D points on a second plane and does not solve the process of mapping from 2D points to 3D points. With this aim, Chen demonstrated that a 3D coordinate system could be added to the laser plane, such that points on the laser plane expand naturally to 3D coordinates by adding a third component equal to zero to the 2D coordinates. Figure 4.5, formerly shown in chapter 2 and reproduced here for convenience,
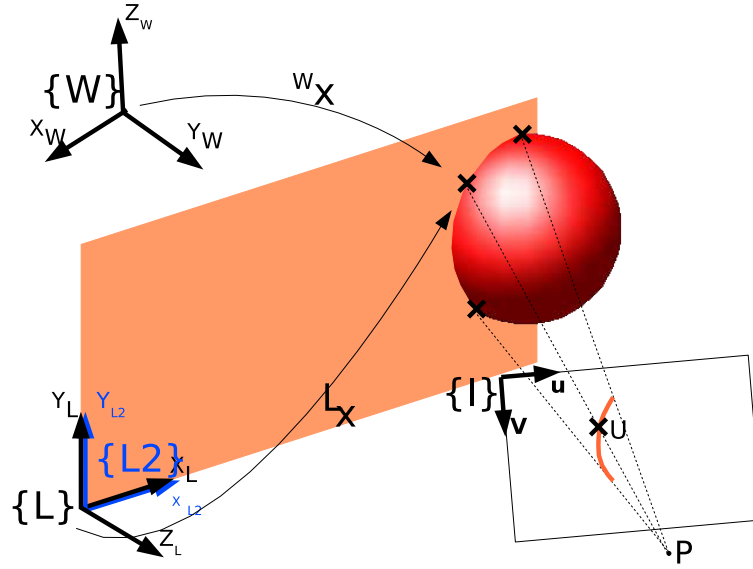
Figure 4.5: Geometric scheme of Chen & Kak's method.

shows a scheme of the above discussion, where $\{\mathbf{W}\}$ is the world or reference coordinate system, $\{\mathbf{I}\}$ is the image coordinate system, with units expressed in pixels, and $\{\mathbf{L}\}$ is the laser coordinate system. $\{\mathbf{L2}\}$ is a bi-dimensional coordinate system, where the x and y coincide with the x and y axis of $\{\mathbf{L}\}$. As shown in equation 4.17, the geometry of the whole system can be modelled by a 4x3 transformation $^W T_I$ with 11 degrees of freedom. This transformation allows the points on the laser plane to be mapped to 3D coordinates, with respect to $\{\mathbf{W}\}$.

$$
\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} =^W T_L \cdot ^L T_{L2} \cdot ^{L2} H_I \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} =^W T_L \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot
$$

$$
\cdot \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \\ t_{41} & t_{42} & t_{43} \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} =^W T_I \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}
$$

$$(4.17)$$

Reid [Rei96] used the projective approach in order to calibrate the Oxford/NEL range finder on the Oxford AGV. In this work, and for the purpose of robot navigation, it was shown that only 8 degrees of freedom were required, instead of the 11 formerly stated. The Oxford/NEL range finder is

a modification of the system proposed by Chen. It also consists of a camera and a laser emitter, but both the image and the laser stripe are reflected on a rotating mirror, so additional geometry is necessarily considered in the system modelling. Both of the previous works, required the pose of the scanner (camera-laser set) to be known precisely with respect to an arbitrary coordinate system, which makes reconstruction dependent on the values of the position controllers. Hence, the scanner has to be mounted on a robot arm, as Chen suggested, or on a mobile robot. In both cases, significant inaccuracies may arise due to the positioning mechanics. Huynh [Huy97] adopted the projective model in order to calibrate a fringe projection range finder. In this work, a different 4x3 transformation matrix is obtained for every fringe. The whole set of fringes is projected at the same time, so every light plane must be uniquely identified in order to set its relationship to the corresponding 4x3 transform. Huynh proposed a method for obtaining point to point correspondences from known 3D points, based on the invariance of the cross-ratio. The 3D data for comparison were obtained using the SHAPE [AN87] scanner. Since a commercial scanner was used, the spatial point resolution is limited by the number of fringes this scanner is able to project. Improvement in performance can only be evaluated in terms of accuracy, since there is no possibility of measuring how fast or how robust it is in the presence of noise. Jokinen [Jok99] used the projective relationship between a laser plane and the image plane as a starting point for refining a laser stripe scanner calibration.

4.5 A new method for obtaining point correspondences

Any set of four non-aligned points A, B, C, D on the plane can be joined pairwise by six distinct lines. This figure is called the Complete Quadrangle and exhibits several useful properties in computer vision applications [MT96] . In this work, 6 of the points defined by this figure are used in order to generate 3D points for calibration and their correspondences on the image plane.

4.5.1 Points on the upper and lower edges

Figure 4.6 shows how the four lines defined by point $\mathbf{F}$ and points A, P_A, B, and G respectively, configure a pencil of lines. Hence it is straightforward to obtain their cross-ratio, provided that point P_A is known. However, P_A is an unknown world 3D point, with only the line that contains it being known,
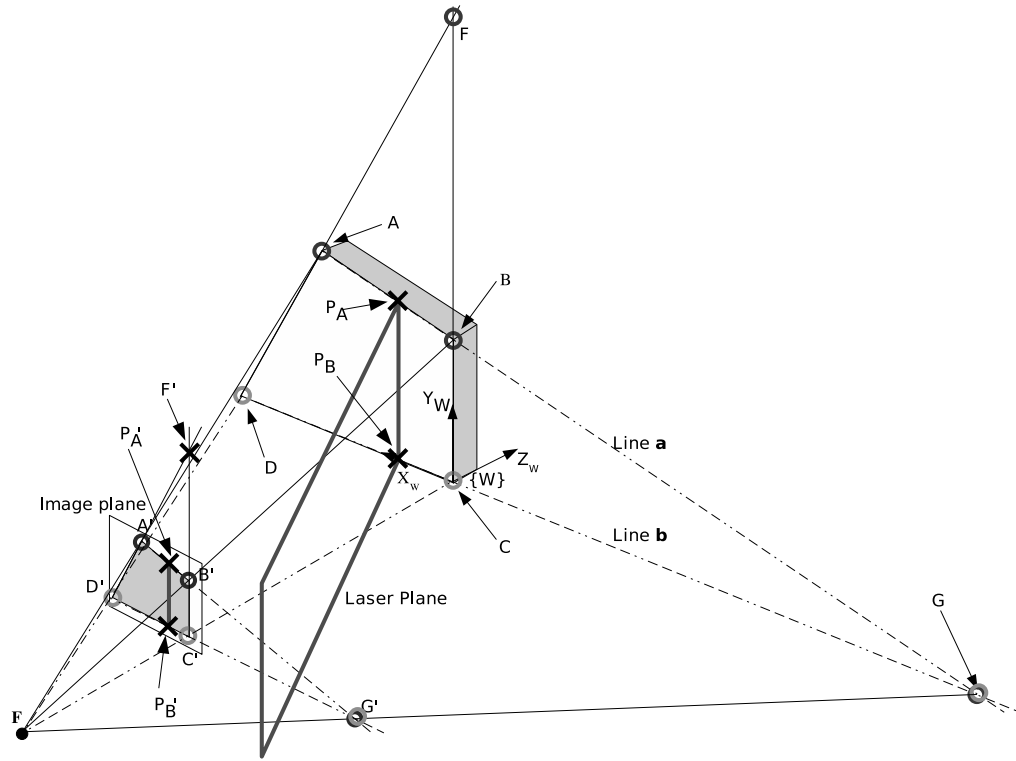
Figure 4.6: The cross-ratio and the complete quadrangle.

that is, only points A, B, and G are known. Nevertheless, if we look at the image, it can be seen that points A', P'_A, B' and G' are all known. Since the geometric relation between points A to G and points A' to G' is projective, the cross-ratio can be computed with points A' to G' and its value can be used for computing the coordinates of point P_A. The same procedure can be adopted between points D to G and D' to G'. By repeating the process for different positions and/or orientations of the laser stripe, it can be seen that a set of correspondences $P_A \Leftrightarrow P'_A$ and $P_B \Leftrightarrow P'_B$ can be established with high accuracy.

4.5.2 Points on the laser stripe

The estimation of the 4x3 transformation ^WT_I, which maps points on the image plane (^IP) to points on the laser plane (^WP), is performed from point

to point correspondences. In subsection 4.5.1, we explained a method for obtaining the intersection points of the laser plane with the upper and lower lines of the calibration target. However, only two points define the laser stripe for each depth, as shown in figure 4.7a. It is reasonable to think, however, that the more points per stripe are obtained, the better the parameter estimation of ^WT_I will be. Looking back to figure 4.6, it is clear that the cross-ratio of points F', A', D' and an arbitrary 2D point P'_L between A' and D' can be obtained. Since points F, A and D are known, the previously computed cross-ratio can be used for calculating the coordinates of the 3D point P_L corresponding to P'_L. This process is depicted in figure 4.7b, where the points on the line $A'D'$ are called P'_L. It is clear that a pencil of lines L'_T can be defined between point G' and points P'_L, and that these lines define a set of intersection points P'_R on the line $B'C'$. Now, for each of the lines of the pencil L'_T, the intersection with the laser stripe P'_S can be obtained, and hence, the cross-ratio between point G', P'_R, P'_S and P'_L can be calculated. Finally, since the 3D points G, P_R and P_L are known, the value of the cross-ratio can be used in order to obtain the set of 3D points P_S laying on the laser stripe.

4.5.3   Validation

In order to validate the quality of the 3D point generation, a plane equation has been fitted using TLS, to the points obtained for each laser plane position, moving the laser stripe linearly, without changing its orientation during the scan. Table 4.1 shows the orientation of the laser plane by using the angles $\alpha = 161.30^o$, $\beta = 90.07^o$ and $\gamma = 71.30^o$, of the direction angles of the plane's normal vector. The position error of the plane is shown as well, in the form of $\Delta X_w = 4$mm, which is the separation between laser positions. In order to see how Gaussian noise influences the point generation, extra noise with $\sigma=1$ and $\sigma=2$ pixels has been added, and the results of the plane fitting are displayed in table 4.1. In addition, the same procedure has been adopted for simulation purposes. It is worth noting that ΔX_w has been chosen to be 4mm in our position controller, however, an offset of 40μm and an error of 45μm are introduced by the mechanics and the controller accuracy, in the case of added $\sigma=0$. As can be seen in table 4.1, the plane orientation is very stable for low noise values, and does not perform badly for very high noise levels. These values show that the point generation algorithm is a good choice for accurate point correspondence generation, since the simulated values obtained are one order of magnitude lower than the ones obtained with real images.
A similar experiment has been arranged, using angular motion, for scanning the laser stripe over the scene. In this case, a plane equation has been fitted

(a)

(b)
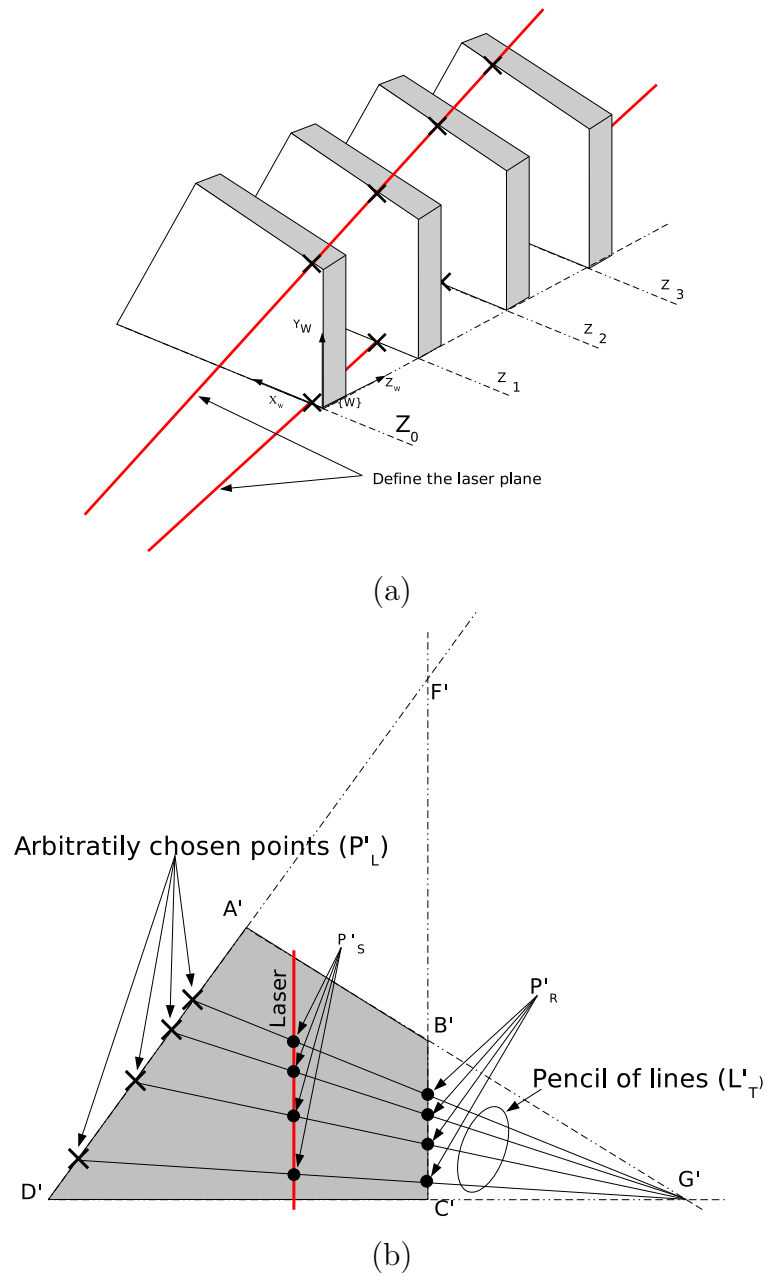
Figure 4.7: The laser plane defined by upper and lower points (a) and Point generation (b).

Table 4.1: Fit of a plane equation to each of the laser plane positions.

Eq.	Real Image error (3σ)			Simulated error (3σ)
	added $\sigma=0$	added $\sigma=1$	added $\sigma=2$	
α	$\pm0.0033^o$	±0.0385	±0.1418	±0.00074
β	$\pm0.0092^o$	±0.2081	±0.1128	±0.00064
γ	$\pm0.0096^o$	±0.1140	±0.4169	±0.00073
ΔX_w	±0.0452 mm	±0.4154	±0.7420	$\pm1.12851e\text{-}06$

for every laser plane orientation and the algebraic distance between each 3D point to the plane has been computed. Then, these data have been used to compute the average distance and the coplanarity error, taken as the standard deviation of the measurements. In addition, the direction cosines and the direction angles have been computed. The results are reported in table 4.2, where it can be observed that the average algebraic distance of the points is considerably small, although for small scanning angles (α or β) a ten times higher bias is observed in comparison to higher scanning angles. Figure 4.8 shows graphically how the errors of coplanarity are more important as the scanning angle increases, which is the opposite behaviour of the average distance. Moreover, figure 4.9 shows how the direction cosines and the angles associated exhibit a linear variation, as expected. In addition, the planes' normal vector y component does not change, since the scanning rotation is performed about L_Y, as opposite to the orientation of its x and z component.

4.6 Calibration

Once the point correspondences have been identified for each stripe position, the parameters of the 2D to 3D mapping must be estimated. Equation 4.18 shows the relationship between points on the image plane $^I[u, v, 1]^T$ with 3D points $^W[sX, sY, sZ, s]^T$ on the laser plane.

$$
\begin{bmatrix} sX \\ sY \\ sZ \\ s \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \\ t_{41} & t_{42} & t_{43} \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{4.18}
$$

Obviously, the parameters t_{11} to t_{34} should be estimated as precisely as possible, in order to maximise the accuracy of the reconstruction. According to
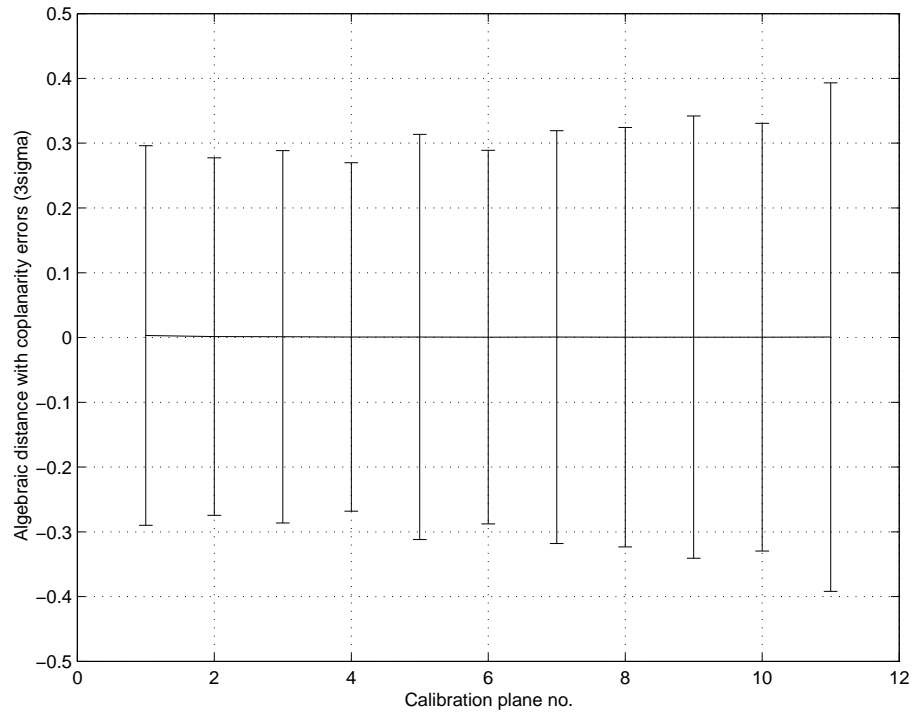
Figure 4.8: Coplanarity errors for each calibration laser plane.

Table 4.2: Coplanarity of the generated planes for angular scanning as a function of the direction angles.

Direction angles			Algebraic distance	Coplanarity error
α	β	γ	$\overline{\mathbf{d}}$	3σ
16.4772	87.9413	73.6591	0.0031	±0.2931
17.5256	87.9288	72.6050	0.0015	±0.2759
18.5703	87.9275	71.5541	0.0011	±0.2874
19.5749	87.9388	70.5427	0.0007	±0.2690
20.6371	87.9306	69.4763	0.0008	±0.3128
21.5901	87.9235	68.5200	0.0006	±0.2883
22.6191	87.9446	67.4848	0.0006	±0.3187
23.5905	87.9601	66.5085	0.0006	±0.3238
24.5523	87.9550	65.5443	0.0005	±0.3415
25.5914	87.9551	64.5023	0.0005	±0.3303
26.5615	87.9572	63.5295	0.0006	±0.3926
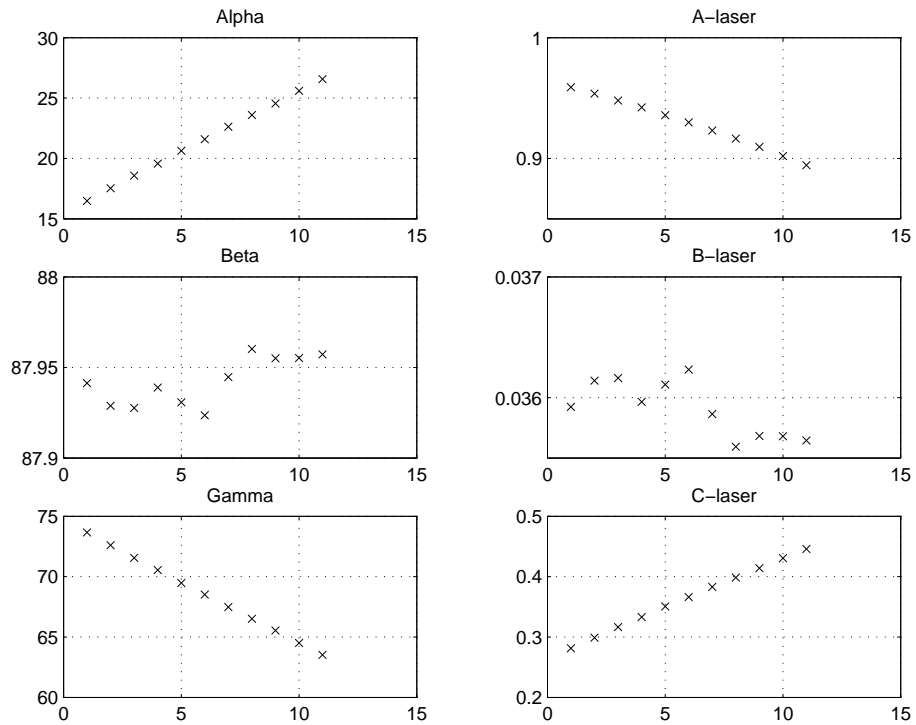
Figure 4.9: Direction cosines and their associated angles for each calibration laser plane.

equation 4.18, the expressions for sX, sY, sZ and s are obtained and shown in equation 4.19.

$$\left.\begin{array}{r} sX = t_{11} \cdot u + t_{12} \cdot v + t_{13} \\ sY = t_{21} \cdot u + t_{22} \cdot v + t_{23} \\ sZ = t_{31} \cdot u + t_{32} \cdot v + t_{33} \\ s = t_{41} \cdot u + t_{42} \cdot v + t_{43} \end{array}\right\} \quad (4.19)$$

Arranging the terms and grouping, a homogeneous system of three equations with 12 unknowns (t_{11} to t_{43}) is obtained, as shown in equation 4.20.

$$\left.\begin{array}{r} t_{11} \cdot u + t_{12} \cdot v + t_{13} - t_{41} \cdot u \cdot X - t_{42} \cdot v \cdot X - t_{43} \cdot X = 0 \\ t_{21} \cdot u + t_{22} \cdot v + t_{23} - t_{41} \cdot u \cdot Y - t_{42} \cdot v \cdot Y - t_{43} \cdot Y = 0 \\ t_{31} \cdot u + t_{32} \cdot v + t_{33} - t_{41} \cdot u \cdot Z - t_{42} \cdot v \cdot Z - t_{43} \cdot Z = 0 \end{array}\right\} \quad (4.20)$$

It can be seen that one single point correspondence contributes with 3 equations in 12 unknowns, with only their ratio being significant. Hence, at least 4 non-collinear points are needed in order to find the 12 parameters. However, due to the presence of noise in the measurement, it is reasonable to transform the problem into a least-squares-type parameter estimation problem, by using more than four noisy point correspondences. Rewriting equation 4.20 in matrix form, equation 4.21 is obtained, which is the expression of a parameter estimation problem, that can be solved by computing the vector θ that minimises some cost function of the matrix equation $A \cdot \theta = 0$, where $\theta = [t_{11}, t_{12}, t_{13}, t_{21}, \ldots, t_{43}]^T$.

A good estimation using TLS techniques can be found by computing the eigenvector corresponding to the smallest eigenvalue of matrix $A^T \cdot A$. In addition, the iterative solution proposed by Van den Hengel [dHBCG00] , the Fundamental Numerical Scheme (FNS), has been tested for estimating the $^W T_I$ parameters. The estimation accuracy of FNS is comparable to that of the Levenberg-Marquardt algorithm, but the rate of convergence is faster. However, FNS requires a good approximation of the covariance matrix for each data sample, which was not available in our experiments. In this case,

FNS performed like TLS, but notably slower.

$$
\begin{bmatrix}
\vdots & \vdots & \vdots & \vdots & & & & & & \vdots & & \\
u_i & v_i & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -u_i \cdot X_i & -v_i \cdot X_i & -X_i \\
0 & 0 & 0 & u_i & v_i & 1 & 0 & 0 & 0 & -u_i \cdot Y_i & -v_i \cdot Y_i & -Y_i \\
0 & 0 & 0 & 0 & 0 & 0 & u_i & v_i & 1 & -u_i \cdot Z_i & -v_i \cdot Z_i & -Z_i \\
\vdots & \vdots & \vdots & \vdots & & & & & & \vdots & &
\end{bmatrix}
\cdot
\begin{bmatrix}
t_{11} \\ t_{12} \\ t_{13} \\ t_{21} \\ t_{22} \\ t_{23} \\ t_{31} \\ t_{32} \\ t_{33} \\ t_{41} \\ t_{42} \\ t_{43}
\end{bmatrix}
=
\begin{bmatrix}
\vdots \\ 0 \\ 0 \\ 0 \\ \vdots
\end{bmatrix}
$$

$$(4.21)$$

With the parameter estimation for each stripe position, a set of $^{W}T_I$ is obtained. In the case of linear scanning, a line can be fitted to the set of parameters corresponding to each matrix component. The case of angular scanning has not been tested in this work, but simulation results show that, instead of a line, a 2nd order function can be fitted. Figure 4.10 shows the values of the set of $^{W}T_I$ and the lines which have been fitted to them. Note that the parameters have been scaled so that $^{W}T_{I34}$ is 1, in order to facilitate the comparison between the matrices.

4.7   Summary

A new calibration method based on the application of projective geometry has been explained. It has been built upon the previous contributions of Chen [CK87], Reid [Rei96] and Huynh [Huy97]. The main contribution of this method is that it allows a laser-in-motion static camera type scanning to be performed, as an improvement of the previous contributions, and this scan can be made either linear or angular. In addition, the number of scanning positions of the laser plane for calibration does not need to be as high as the number of scanning positions, just a few positions are sufficient to get a good calibration. Nevertheless, the resolution for scanning a surface can be arbitrarily chosen, hence offering the possibility of obtaining a very dense cloud of points. Another advantage of this method is that there is no need to obtain the physical geometry of the system, since the whole geometry is modelled projectively, by means of a set of 4x3 matrices. Moreover, the absolute motion of the laser plane does not need to be known but only the initial position and the increments should be under control. The results
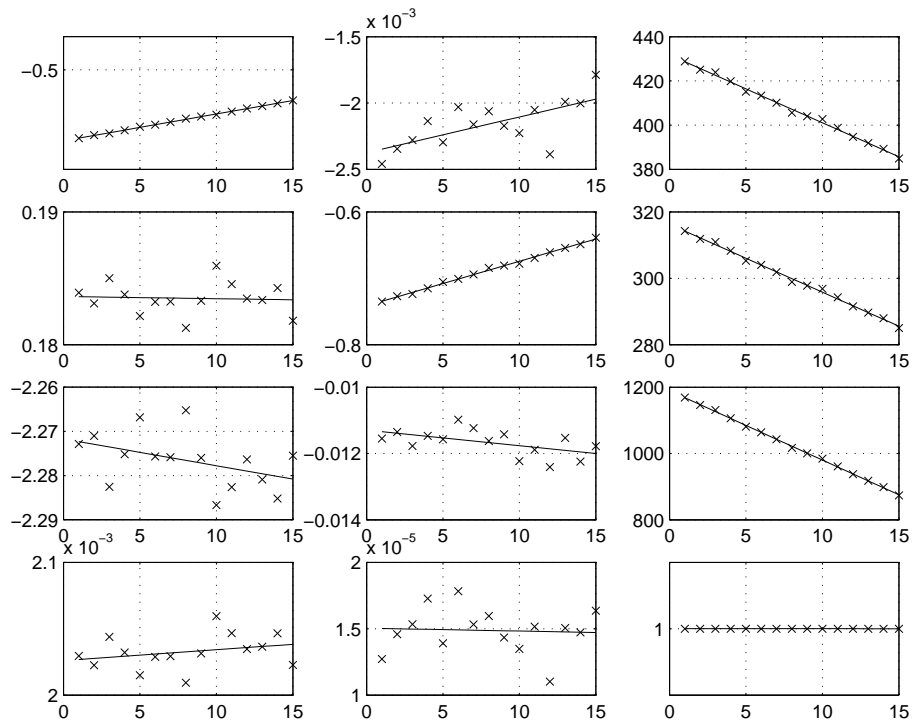
Figure 4.10: Plots of the values of $^{W}T_{I}$ and the lines fitted (linear scanning).
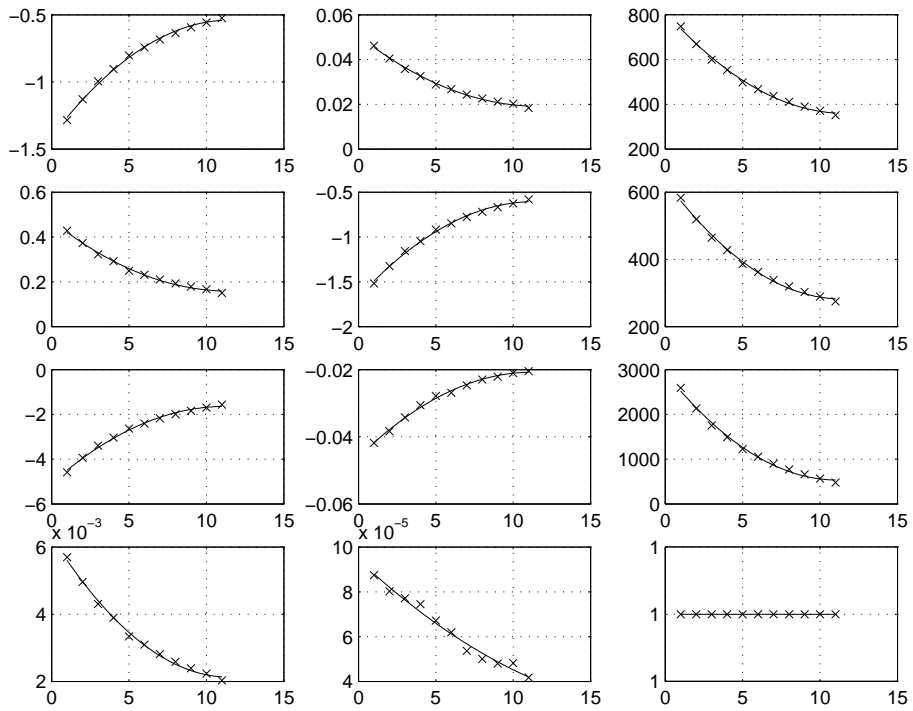
Figure 4.11: Plots of the values of $^{W}T_{I}$ and the curves fitted (angular scanning).

which have been obtained are very encouraging, and they have been reported in chapter 5.

Chapter 5

Experimental results

This chapter reports on the results obtained during the test of both the peak detector and the calibration method, explained in chapters 3 and 4 respectively. Since the validation of the 3D point generation for obtaining point correspondences to feed the calibration algorithm has been reported in chapter 4, only reconstruction performance is described here. To this end, the reconstruction of a known cylinder has been done, and its diameter has been used as a benchmark for obtaining the error of reconstruction. The tests undertaken with the peak detector are done under two extreme conditions of a) a lot of light scattering impinging on a translucid material, and b) a very low light level impinging on an opaque lambertian surface.

5.1 Test bench

In order to carry out with experiments for validating the proposed algorithms, a test bench has been used composed of an optical table, with a grid of 6mm screw holes separated 25mm$\pm$4μm to each other. On this table, a linear stage holding the semiconductor laser emitter has been mounted. The linear stage resolution is 2μm. In addition, a grey level CCD camera with standard CCIR pixel resolution (752$\times$582 pixels) has been used. The laser emitter is a semiconductor diode which outputs light at 1mW peak power with λ=650nm. In addition, a cylindrical lens is provided with the laser housing so that it is capable of projecting a plane of light with an 85° aperture and 1mm width, and a bandpass optical filter with BW=10nm has been used for segmenting the laser stripe from the rest of the scene. The tests have been undertaken using a 25mm C-mount lensmounted on camera with a frame-grabber which provides standard VGA (640x480 pixels) resoltion and 8 bit grey level. The distance of the object under consideration is about 1200mm, while the laser

incidence angle with respect to the optical axis of the camera is approximately 18.5^o. Figure 5.1a shows a photograph of the test bench with the material used, mounted on it, and figure 5.1b depicts the scheme of the scanner setup, with the measuring distance, and the laser angle of projection.

5.2   Experiments with the peak detector

In order to evaluate the effect on the peak estimation, two experiments have been arranged. The former consists on evaluating the behaviour of the proposed method on a mate surface, for different S/N values, comparing with 5 other peak estimators. The performance of these peak estimators has been previously studied by [FN91], and are known as Gaussian approximation (**GA**), Centre of mass (**CM**), Linear approximation (**LA**), Blais and Rioux detector (**BR**) and Parabolic estimator (**PE**). The thorough description of these methods falls beyond the scope of this chapter, however, table 5.1 enumerates their mathematical equations, and a comparison of their performances was published by Fisher [FN91]. The BR detector, uses a signal $g(n)$ for estimating the peak position. This $g(n)$ is the filtered and derived row signals, using the filter proposed in the work of Blais and Rioux [BR86].
In the second experiment, a piece of translucid white plastic has been partially painted in a mate, white colour. The thickness of the paint layer is $60\mu m$ in average. This experiment is similar to that performed by [GRB$^+$01]. The filter coefficients have been previously estimated using the Matlab *fda-tool* function.
In the first experiment, the image with the lowest S/N has been chosen for computing the filter coefficients. In the second experiment the most noisy side of the image is the one corresponding to the response of a translucid surface, hence this side of the image has been used for obtaining the filter coefficients. In both cases, the stop-band has been chosen to have an attenuation of 80dB. The cut-off frequency and the width of the transition band have been selected after the study of the row signals in the frequency domain. In the first experiment, the filter order has been 56 while for the second experiment the filter order has been selected to be 296. These results are consistent with the fact that the laser impinging on the translucid surface show a much higher noise level, and that the stripe is significantly wider in the second case than in the first one.
In the first experiment, the 6 peak estimators have been applied to the whole stripe image, computing the standard deviation of the peaks on each image, for 4 S/N levels. This results are summarised in table 5.2. In this table, the proposed method [FSCP04] has been identified with the acronym "**PM**". As

(a)

(b)
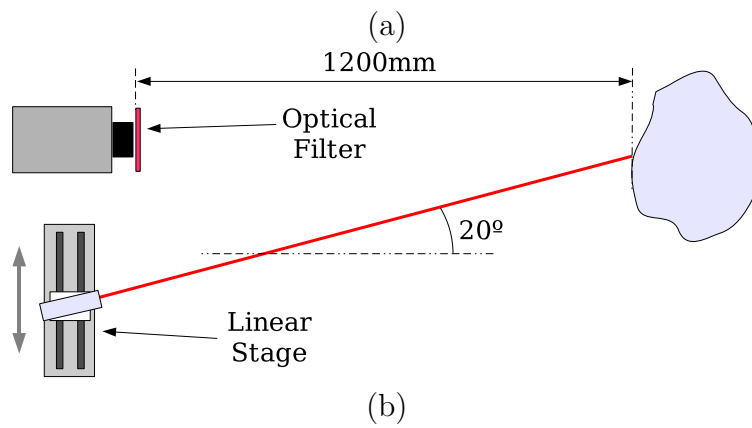
Figure 5.1: A picture of the test bench (a) and the lab scanner.

can be observed, PM keeps a lower and more constant value of σ compared to the other methods for any S/N value, even when it is considerably low. In addition, figure 5.2, shows the results for the lowest S/N graphically. As can be seen, under very low light conditions, the lowest performance is exhibited by BR, as table 5.2 reports numerically.

In the second experiment, the stripe is considerably wider than in the former one, with a large saturation width (i.e. each row peak is far from fitting a pure Gaussian shape). This prevented us to use the 6 methods, since GA, LA and PE can not be applied without modifications in the algorithms. Instead, we have used only PM, BR and CM algorithms, computing the standard deviation in both sides -opaque and translucid- of the surface for two S/N conditions. It is worth noting that speckle noise is much more significant under these contitions than thermal noise. This makes the S/N levels to be lower for high light power than low light power. In addition, considering the peak estimation on a translucid material, this effect is magnified due to the light scattering. Furthermore, a bias is observed when estimating the peak on the translucid material. Table 5.3 compiles the peak estimation data in terms of standard deviation and mean values. It is worth observing that for opaque surfaces, PM shows better results (in terms of σ) for low light power than for high power. This is due to the wider image of the stripe for high light power. The peak estimation for translucid surfaces, however, is better for high than low light power. CM is the best option under these conditions, but it has been shown previously that CM is not a good choice when the stripe is narrow. In addition, CM is very sensitive to noise, in the sense that it experiences offsets due to relatively small perturbations. Comparing PM and BR, using filters of the same order but different coefficients, PM performs better in both light power conditions and both types of materials, especially with low S/N. In addition, the logical bias in the measurement of the translucid material ($60\mu m$ of paint thickness) varies in both light power situations, and is more significant using BR than PM, as stated in the column labelled $\triangle \hat{X}$ in table 5.3. PM and BR are much more sensitive to light changes than CM, however, a visual inspection let us observe that $\triangle \hat{X}$ is closer to reality under low light power and better estimated by PM and BR. The results are graphically depicted in figures 5.3, 5.4 and 5.5 for PM, BR and CM estimators respectively.

Table 5.1: Estimator formulae. The $\hat{\delta}$ stand for the subpixel offset. The a,b and c stand for the 3 consecutive pixels of the peak, where b is the maximum in intensity value. The x_o, y_0, x_1, y_1 of PM are the coordinates of two points to which a line is fitted, as explained in chapter 3.

Estim.	Formulae
BR	$\hat{\delta} = \begin{cases} \frac{g(i)}{g(i)-g(i+1)} & f(i+1)>f(i-1) \\ \frac{g(i-1)}{g(i-1)-g(i)} & f(i+1)<f(i-1) \end{cases}$
GA	$\hat{\delta} = \frac{1}{2} \cdot \left(\frac{\ln(a)-\ln(c)}{\ln(a)+\ln(c)-2\cdot\ln(b)} \right)$
CM	$\hat{\delta} = \frac{c-a}{a+b+c}$
LA	$\hat{X} = \begin{cases} x - \frac{a-c}{2(b-a)} & c > a \\ x - \frac{a-c}{2(b-c)} & otherwise \end{cases}$
PE	$\hat{\delta} = \frac{1}{2} \cdot \frac{a-b}{c-2\cdot b+a}$
PM	$\hat{X} = x_o - \frac{y_0 \cdot (x_1 - x_0)}{y_1 - y_0}$

Table 5.2: Values of σ estimating the peak using the 6 methods with 4 S/N levels.

SN ▷	13.34 dB	8.12 dB	4.13 dB	0.92 dB
PM	1.719	1.712	1.728	1.777
BR	1.755	1.799	1.910	4.228
CM	1.744	1.809	1.813	1.892
GA	1.759	1.842	1.816	1.910
LA	2.194	3.088	2.159	1.890
PE	1.792	1.833	1.820	1.889
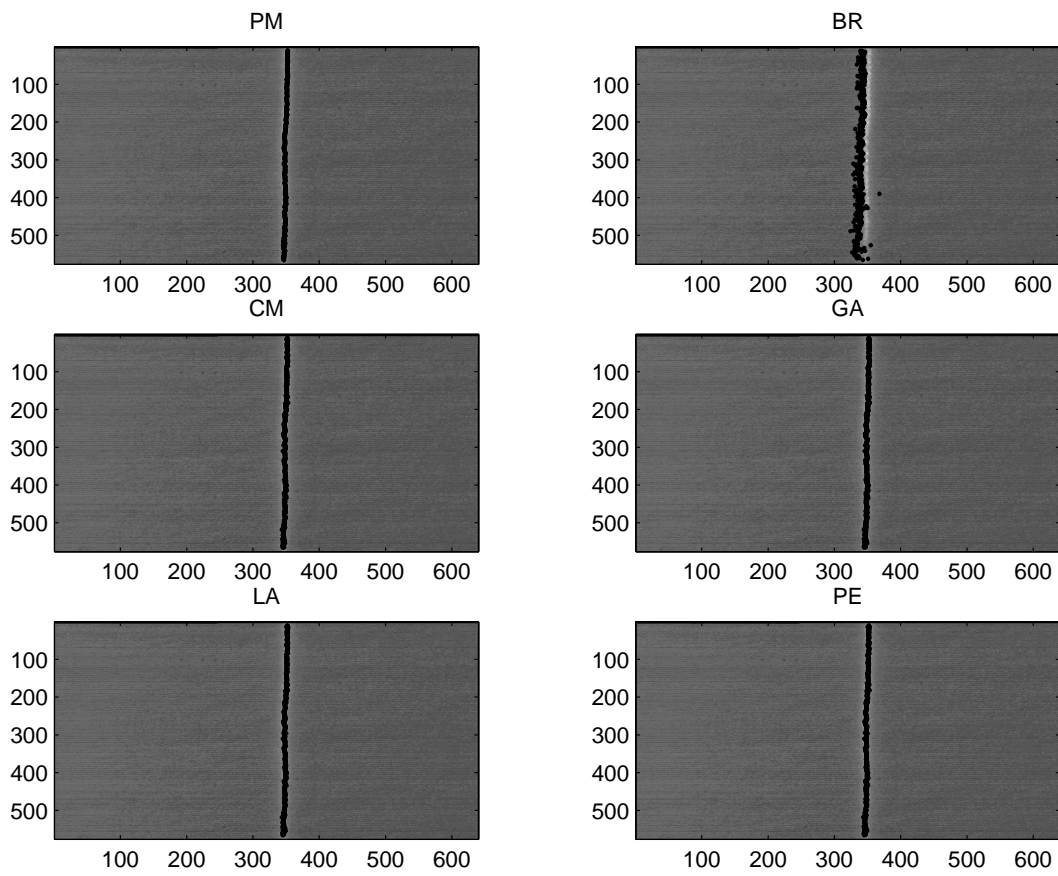
Figure 5.2: Peak estimation with the 6 methods, SN=0.92dB.

Table 5.3: Mean value and σ, estimating the peak on two types of material under different light power conditions (S/N values in dB).

$\triangle$	Opaque (11dB)		Trans. (7 dB)		
Light	$\hat{X}$	σ	$\hat{X}$	σ	$\triangle\hat{X}$
PM	382.841	1.862	338.361	4.258	44.480
BR	380.719	2.314	321.756	4.560	58.963
CM	378.858	1.895	350.054	2.132	28.804
∇	Opaque (21dB)		Trans. (12 dB)		
Light	$\hat{X}$	σ	$\hat{X}$	σ	$\triangle\hat{X}$
PM	383.647	1.639	365.478	6.010	18.168
BR	382.544	1.640	363.043	9.363	19.501
CM	379.732	1.519	355.054	2.687	24.677
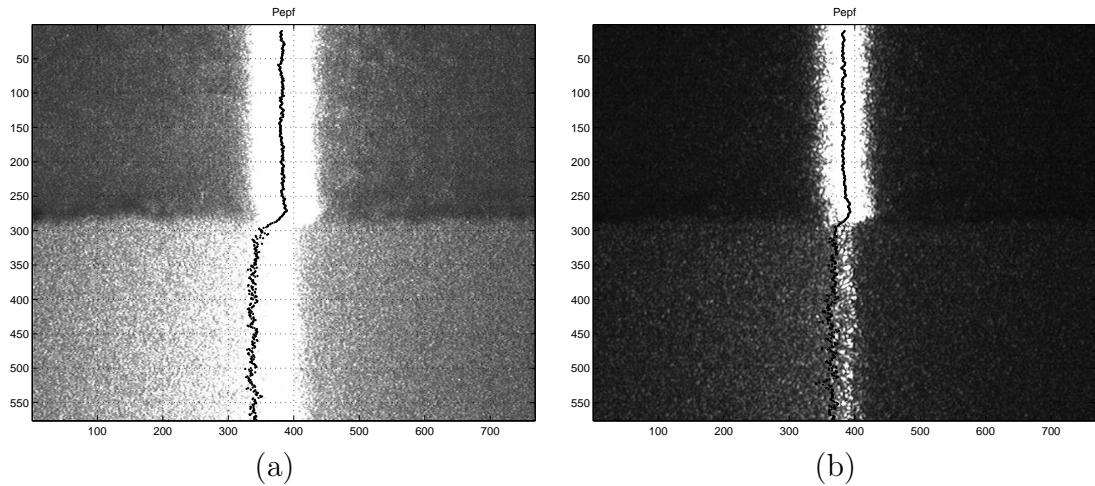
Figure 5.3: Effect of peak detection on a lambertian (up) and translucid (down) surface. Using **PM**.
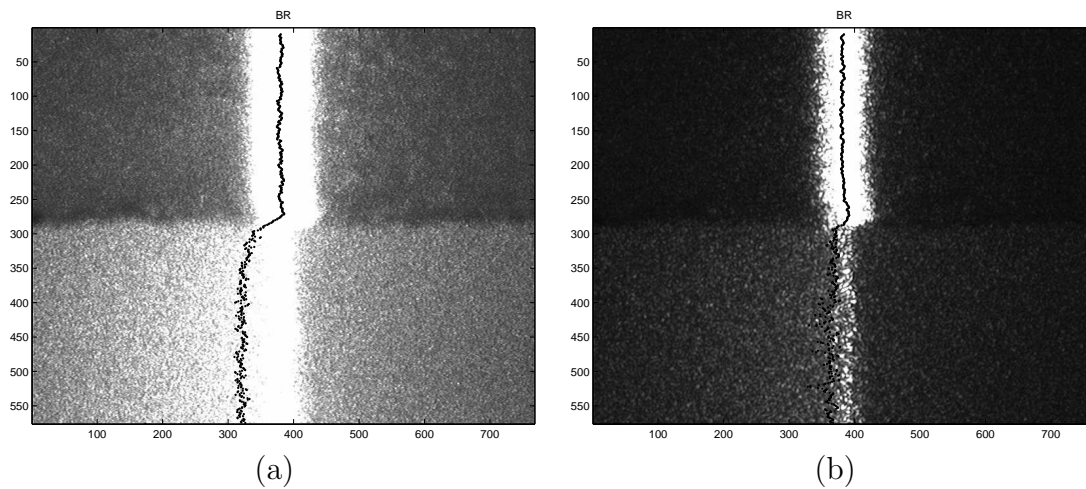
Figure 5.4: Effect of peak detection on a lambertian (up) and translucid (down) surface. Using **BR**.
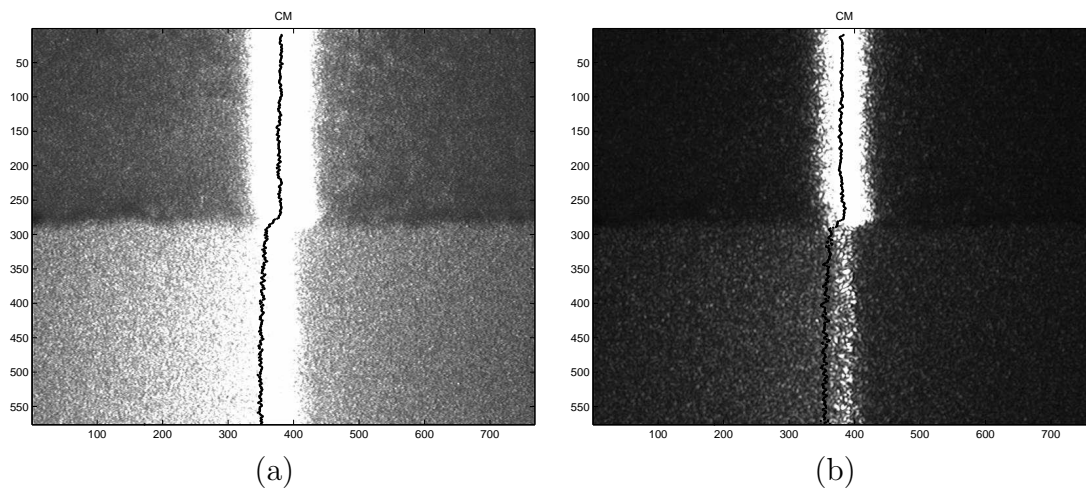
Figure 5.5: Effect of peak detection on a lambertian (up) and translucid (down) surface. Using **CM**.

5.3 Experiments with the scanner calibration

As explained in chapter 4, our method of calibration consists in obtaining a set of 4×3 matrices which map 2D points to 3D points, for each position of the laser stripe. In order to see how good is this method, two experiments have been undertaken. The former consists in reconstructing the 2D points ($^I\hat{P}$) obtained from the calibration images, as explained in section 4.5. In the ideal case of computers having infinite precision and the stripe images being taken in the absence of noise, the difference between the 3D points obtained by this procedure ($^W P$) and the 3D points obtained by the complete quadrangle method ($^W\hat{P}$) is zero. However, computer operations exhibit limited performance due to their limited accuracy. Furthermore, image sensors are subject to electrical noise during the imaging process, and the mechanics used for scanning the laser stripe are subject to position inaccuracies. Hence, due to these diverse noise sources, reconstructing the calibration planes is subject to errors, which are displayed as offsets and discrepancy between $^W\hat{P}$ and $^W P$, as displayed in figure 5.6. The narrower error interval corresponds to the values without added Gaussian noise, the middle and the wider values correspond to added noise of $\sigma=1$ pixel and $\sigma=2$ pixels, respectively. The calibration algorithm has been tested using the scanner setup shown in figure 5.1b.

The second experiment consists in scanning a well known object with simple geometric features. We chose a cylinder for this object with ø=73.30mm, measured with a digital caliper with $\pm 20\mu$m accuracy. The 3D reconstruction of such cylinder is shown in figure 5.7. In order to extract the diameter, a TLS scheme has been used to fit a cylinder to the reconstructed data, and the data has been added Gaussian noise with $\sigma=0.4$ to 2 pixels. The results are summarised in table 5.4, and the measurements have been taken at a distance of 1100mm.

5.4 Qualitative results

Figures 5.9 to 5.12 show a number of digitised statues. They have been taken by the prototype introduced in chapter 4, incorporating the peak detector explained in chapter 3 and it has been calibrated using the method described in chapter 4. Figure 5.8 shows a cloud of points as it outputs the system, whith no post-processing. Figures 5.9 to 5.12 show the corresponding surfaces, which have been warped a mesh of triangles, together with one of the profiles reconstructed for each of them. The bust of Wagner and the horse have been acquired at a distance of 1250 mm, while the laughing sun and
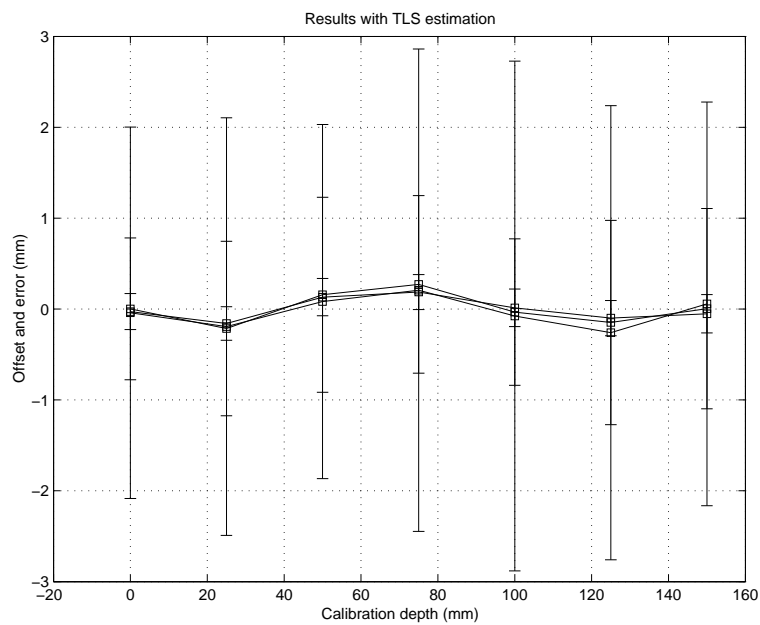
(a)
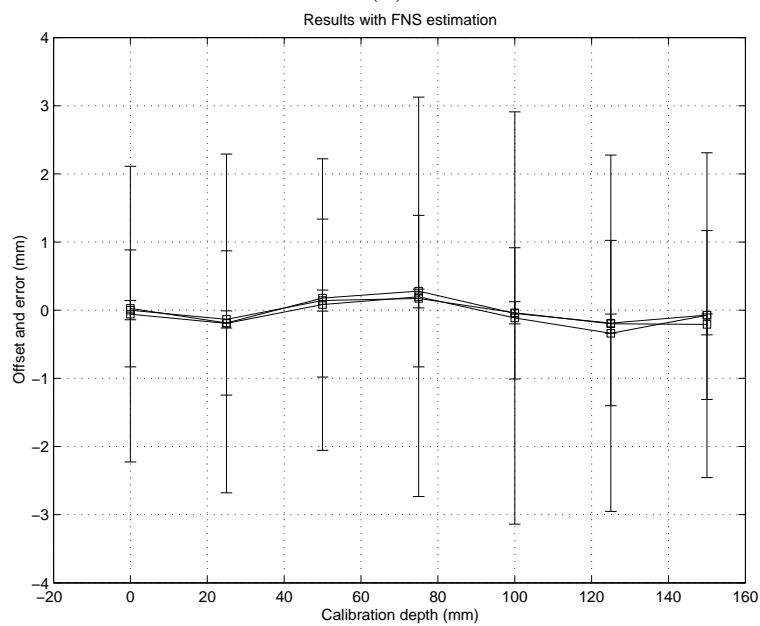
(b)

Figure 5.6: Errors in the reconstruction of the calibration planes using (a) TLS and (b) FNS methods.

Table 5.4: Reconstructed diameter of the fitted cylinder with different noise levels.

Added σ	Diameter ø	% Rel. error
0	73.52	0.3
0.4	73.16	0.2
0.8	72.07	1.7
1	71.14	3
1.4	69.21	6
1.8	66.90	9.6
2	65.38	12.1
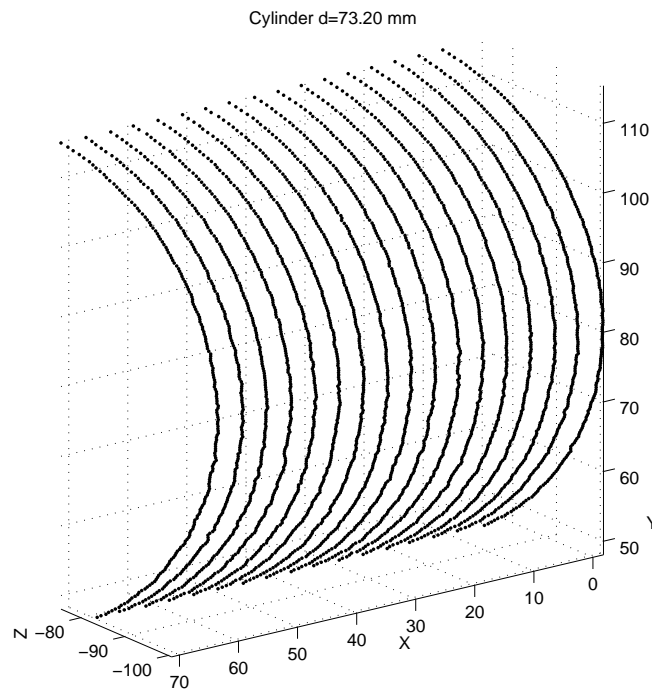
Figure 5.7: Reconstruction of a cylinder (a).

the samurai have been digitised at about 1000 mm.

In order to test the ability to digitise translucid surfaces at great distances, a plastic surface of this type with letters carved on it was arranged. The depth of these letters is 170μm in average, and the surface was digitised at a distance of about 600mm to the lens. As can be seen in figure 5.13, the letters are perfectly readable by comparison with the acompaining photograph. In addition, figure 5.14 shows a reconstruction of a portion of human tissue. The holes' diameter is 50μm. Unfortunately a picture is not available, but when the tests were performed, it was seen that the translucidity was much higher than that of the plastic used in the previous tests, and hence it proves that our peak detector approach performs very well under such adverse conditions.

5.5   Summary

A set of both quantitative and qualitative results have been reported, which gives a comprehensive overview about the systems performance, including both the peak detector algorithm, the calibration algorithm using the complete quadrangle and the reconstruction procedure. The reconstructed surfaces have been obtained as a set of points each, and a mesh of triangles has been interpolated. The triangle meshing and representation has been done using the VTK library, developed by Kitware, Inc. and freely distributed on the web.
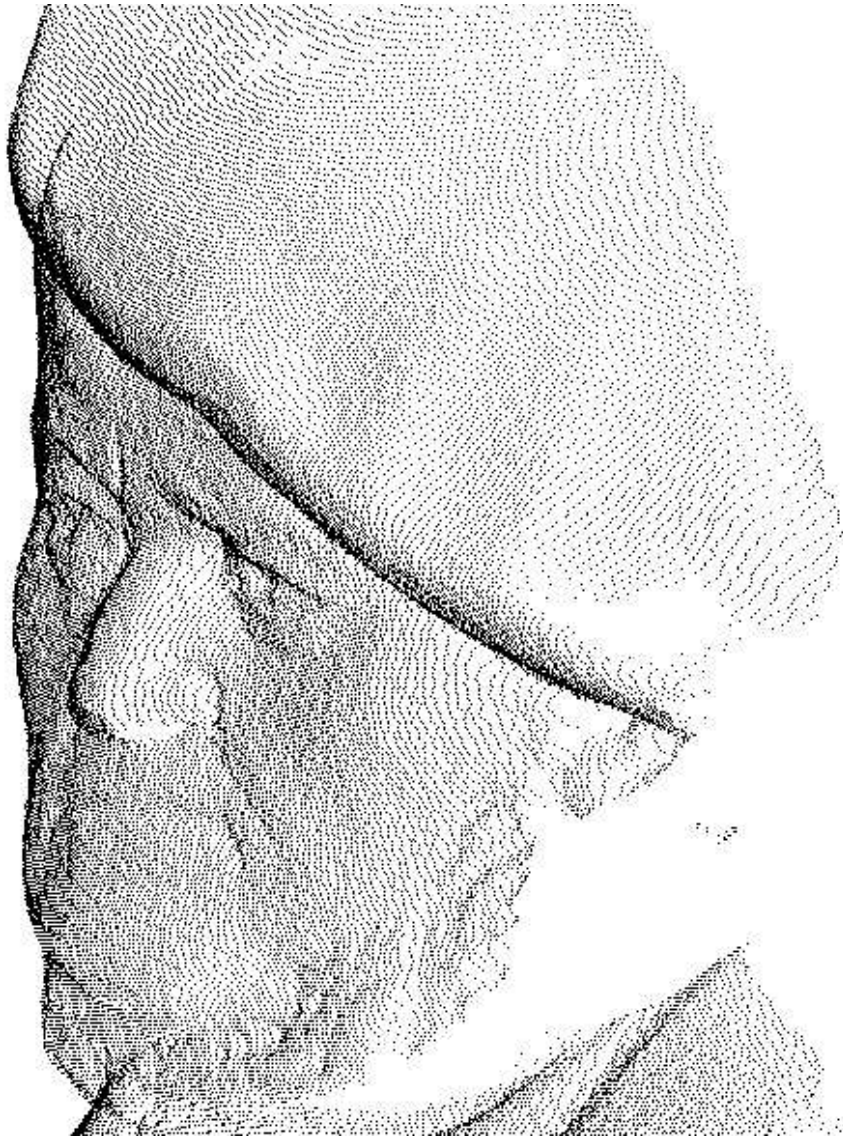
Figure 5.8: A dense cloud of points of a bust of Wagner.

(a)   (b)

(c)
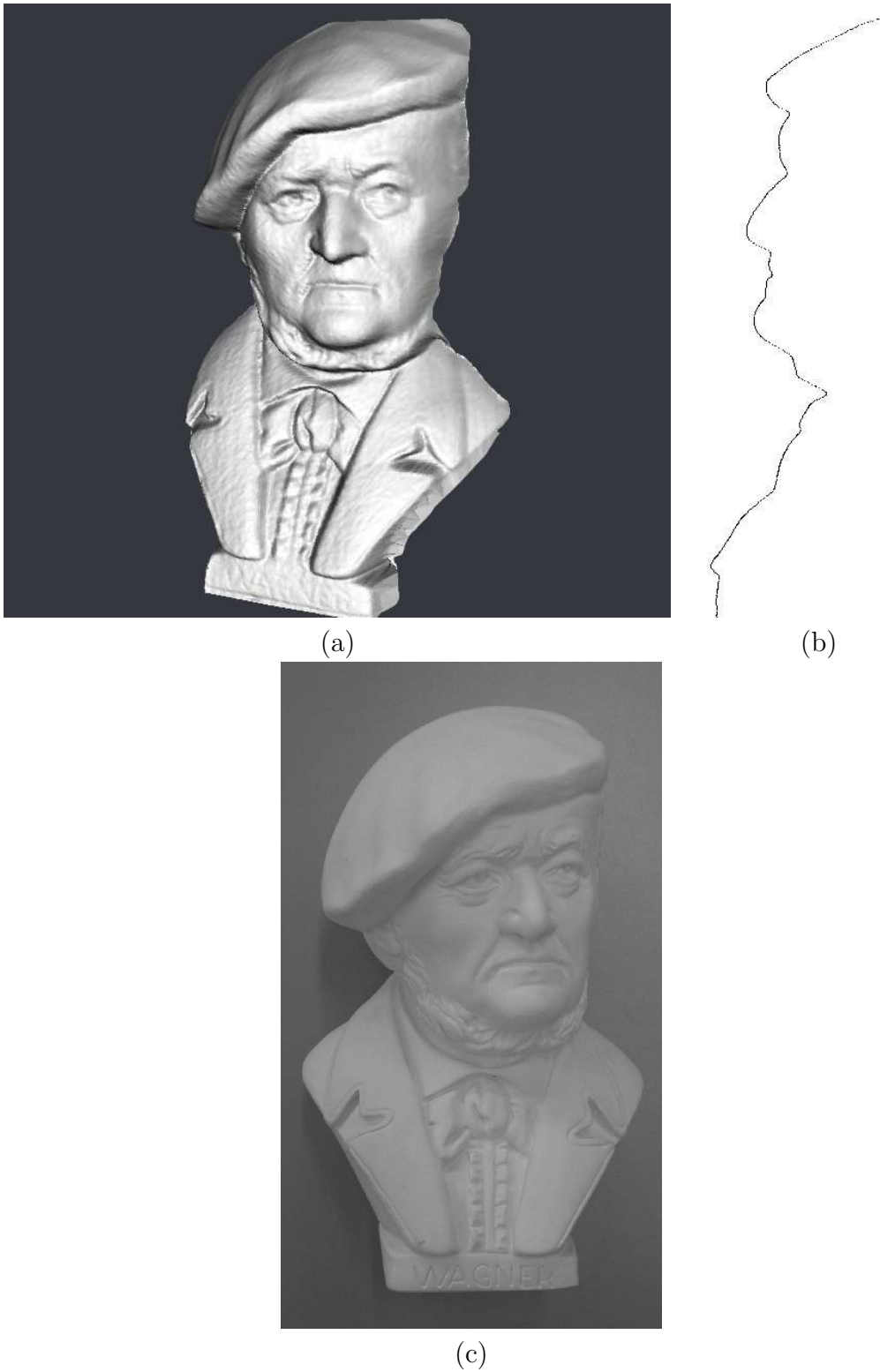
Figure 5.9: The bust of Wagner: interpolated surface (a), the extracted profile (b) and its picture (c).

(a)   (b)

(c)
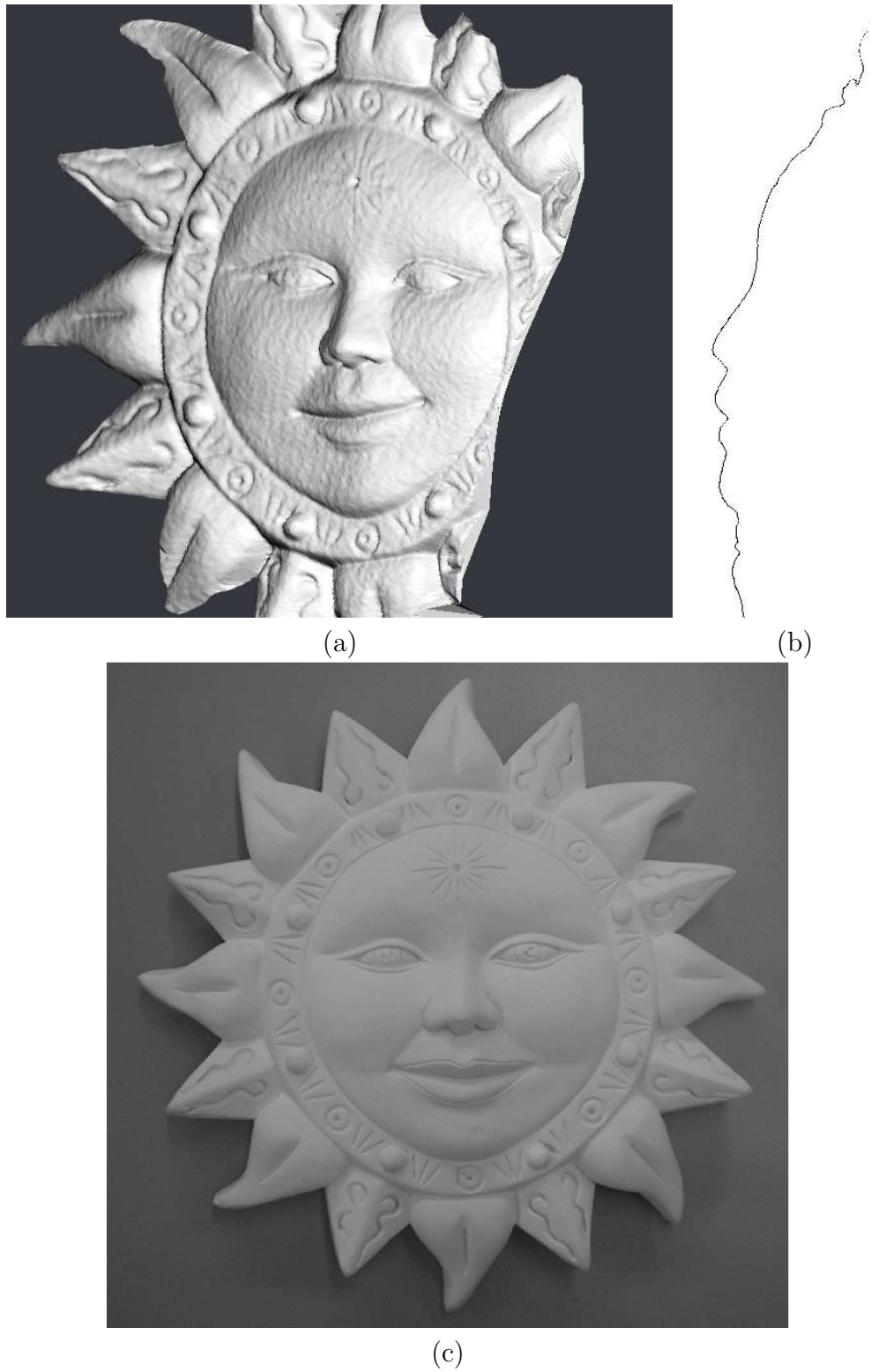
Figure 5.10: A laughing sun: interpolated surface (a), the extracted profile (b) and its picture (c).

(a)   (b)

(c)
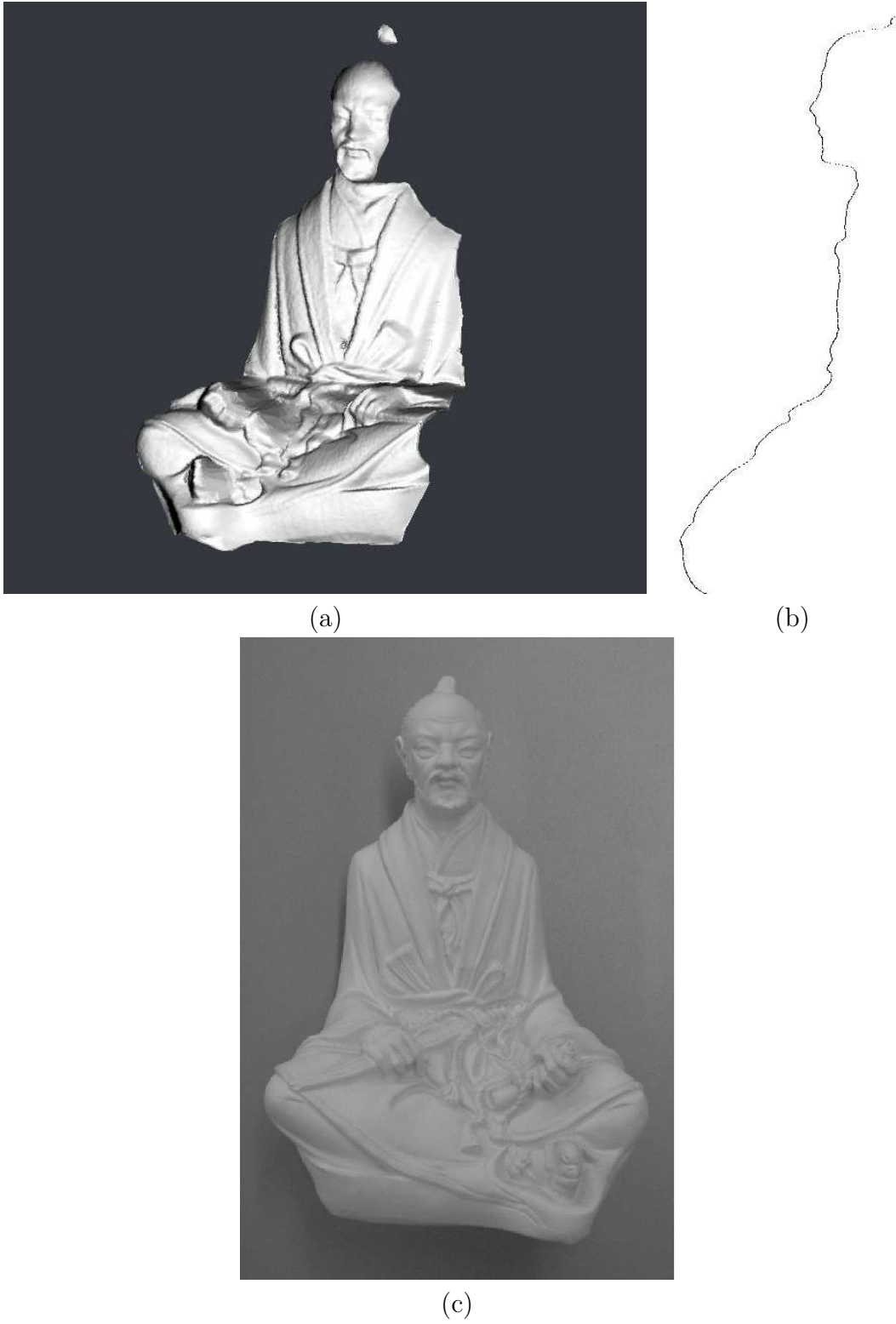
Figure 5.11: A meditative samurai: interpolated surface (a), the extracted profile (b) and (c) its picture.

(a)

(b)

(c)

Figure 5.12: The bust of a horse: Interpolated surface (a) and its picture (b).
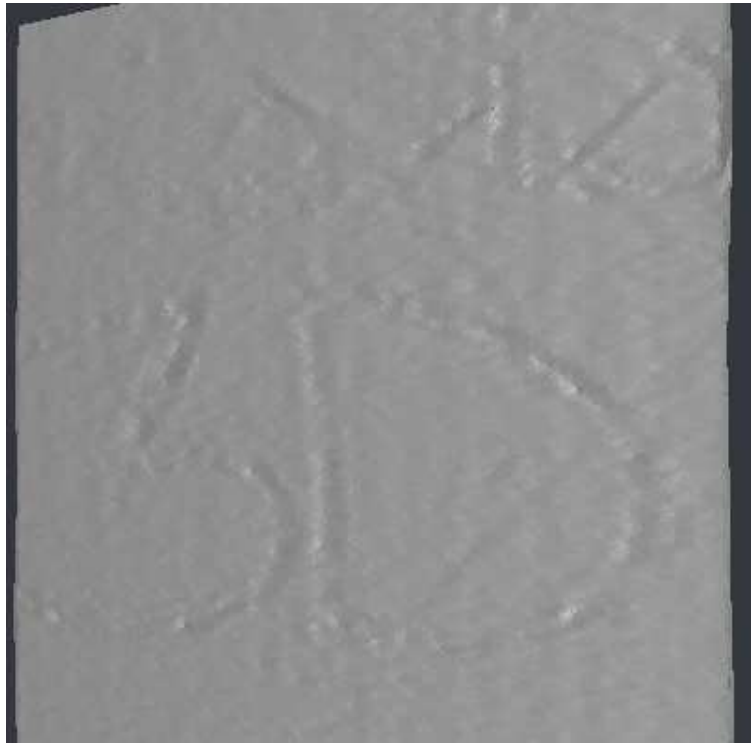
(a)

(b)

Figure 5.13: Reconstruction of a translucid surface with letters carved on it (a) and its picture (b).
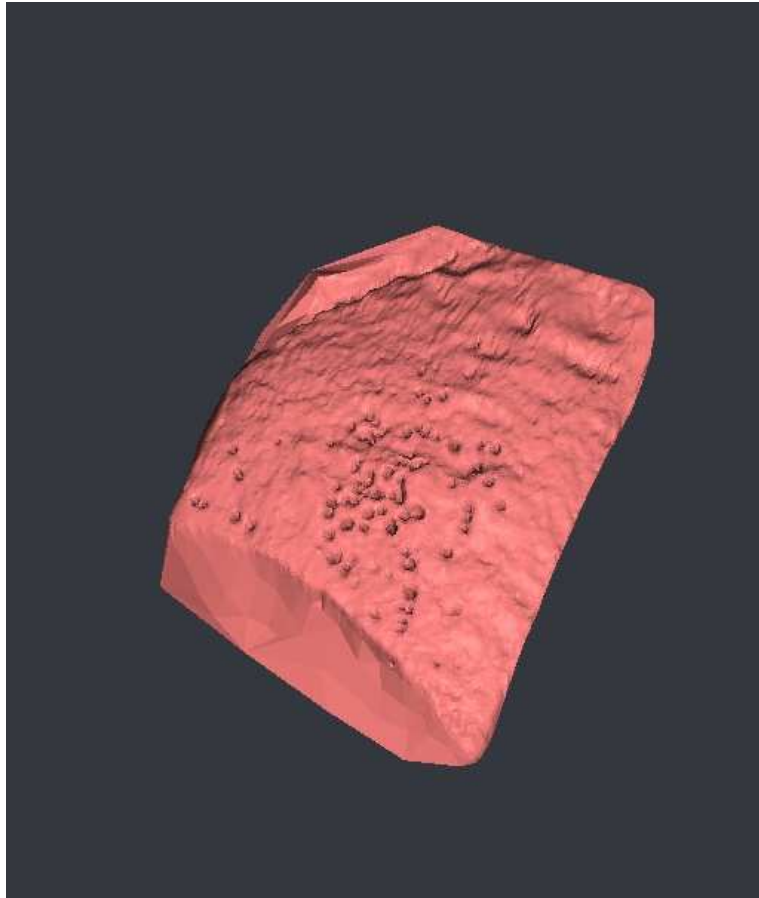
Figure 5.14: Reconstruction of a portion of human tissue.

Chapter 6

Conclusion

The digitisation of 3D surfaces using triangulation laser scanners is still an open problem in many aspects, although there is an extensive literature reporting solutions to some of them. It is a vast field comprising very diverse applications, and it is developing toward new sectors. Even when there is a vast number of research and applications beyond the shape acquisition, such as surface description, surface fitting, polygonal meshing, etc, great issues like accuracy improvement in recontsruction or acquisition speed are always welcomed, specially in the field of robotics and visual servoing. Even when there exist good techniques for improving surface quality post-acquisition, if the process of shape digitisation is improved, a lot of processing time can be saved and a best and faster surface fitting is obtained. Two of the topics which more affect the improvement in acquisition accuracy are: 1) The laser peak detector and 2) the calibration process. This thesis addresses these points by providing a new numerical peak detector, as seen in chapter 3, and a novel algorithm for calibrating 3D stripe range finders projectively, as seen in chapter 4. In addition, chapter 2, provides a representative overview of scanners with emphasis on their calibration. In the following, a formal set of contributions found in this thesis is listed. Then, the conclusions of each contribution are drawn separately. Finally, a series of directions are discussed for improving the results obtained. In addition, a list of publications arising from this work is provided.

6.1   Contributions

All the contributions of this work are enumerated in the following. They are briefly summarised and structured as items. The enumeration has been done considering the chronological path which has been followed along the

years until the new peak detector and the new calibration algorithm where developed.

- An exhaustive review of triangulation-based laser slit scanners. It provides a comprehensive reference table with citations to the authors of each calibration and reconstruction algorithm.

- The implementation and evaluation of the former projective calibration technique for laser stripe range finders. This work [CK87] is the basis upon which one of the contributions has been developed.

- The implementation and evaluation of a prototype of a smart sensor. A commercial photo-diode array has been used for exploring the possibilities of smart sensors.

- The implementation and evaluation of a laser range finder for underwater applications. In this work, although a physical model was used for calibration, the performance of these devices has been evaluated and the particular issues like scattering, absorbtion and extinction of light underwater were studied.

- A new algorithm for the obtention of a robust estimation of the peak position of each row image of a laser stripe. This new technique allows the scanning of translucid materials and its application to 3D scanning under significantly low signal-to-noise ratios.

- A novel algorithm for calibrating a triangulation laser stripe 3D scanner. It is developed on the basis of projective geometry and a new technique for obtaining point to point correspondences has been reported, which obtains the most accurate 3D point corresponding to a given 2D point.

6.2   Conclusions

In the following, conclusions for each contribution are displayed. The order of appearance corresponds to the order they have been explained throughout the thesis, and they chronologically correspond to the papers listed below.

6.2.1   Survey of triangulation laser devices

An exhaustive review of representative laser slit scanning three-dimensional digitiser systems has been reported. A classification table has been proposed

showing the main features for each system, including references to the corresponding authors, expanding the discussion about the different calibration methods.

There are no two identical calibration methods, but a different one exists for every different digitiser. Nevertheless, according to the nature of the hardware involved in each system, accuracy is always strongly dependent on the system calibration as well as the stripe detection accuracy.

The survey contributes a wide range of literature references, where more than 100 journal and conference articles have been reviewed. In addition, emphasis has been put only on triangulation-based devices, which gives the survey a high degree of concrecity. Moreover, the former two applications reported in appendix A, were selected from this study and were specially implemented with the aim of building the basis for developing the rest of the thesis.

6.2.2   Applications

The method of Chen & Kak has been reproduced as a simulation and a working prototype has been implemented in order to obtain a more thorough accuracy evaluation. The real implementation of such system has allowed our research lab to include a digitisation facility for research purposes, as well as experimentation with new calibration algorithms as well as new stripe detection operators.

A new calibration technique for smart sensor-based digitisers has been reported and a working prototype has been implemented using a standard reflex camera housing. Although a simple 5x5 photo-diode imager sensor has been used, it is sufficient for acquisition speed and accuracy evaluation as well as for image acquisition hardware design for range or $2\frac{1}{2}$D images. In addition, it is our purpose to continue with this field of application and a new digitiser based on a similar principle of operation is being envisaged, such that it will bring significant improvements in both the resolution and accuracy of the surface acquisition, as well as a novel technique for sensing and processing light on-chip.

6.2.3   Peak detector

A new method for estimating the peak position of a laser stripe has been reported, and its performance has been compared with other 5 existing methods. The experimental results show that the proposed method yield to better estimations of the peak position, especially when the S/N is very low. When translucid surfaces are scanned, a bias in the peak estimation appears, which is a function of the impinging light power. The computational complexity of

the proposed method is similar to that of BR [BR86] but increases its accuracy in a wider range of stripe light power or non-cooperative surfaces. The strong point of this method is the analysis of the row signals in the frequency domain, which yields to the cut-off frequency and the transition band width for obtaining the coefficients of an optimised filter. In addition, it is clear that speckle constitutes a fundamental limit in the stripe detection, hence it would be of interest to use a non coherent light source, or a speckle-free laser light, which is an emerging technology in laser design for imaging purposes.

6.2.4   Calibration method

A novel and elegant method for obtaining highly accurate 2D to 3D point correspondences for calibration has been proposed. In this work, we used it for the calibration of a laser 3D scanner, although the method is also valid for standard camera calibration. In addition, the projective calibration approach for 3D laser scanners has been extended for use in *scanning laser, static camera*-type systems. One of the most powerful features of the projective approach is that no physical model of the system is necessary, since the whole geometry is contained in a single matrix. This feature has been kept in our proposal, since the whole system geometry is contained in a *set* of matrices. The experiments undertaken in our laboratory facilities exhibits very good results considering the components used for configuring the set-up. It is clear that the most influential issue in this calibration algorithm is the parameter estimation optimisation algorithm. In order to improve the performance, the parameter estimation using both FNS and Levenberg-Marquardt algorithms has been carried out. However, the performance of both FNS and Levenberg-Marquardt is significantly affected by a good estimation of the covariance matrix for each observation. Since the system calibration is done under controlled conditions, the estimation of the covariance matrix can be obtained from successive samples of the same laser position for different lighting conditions. It is the belief of the authors that the accuracy can be improved by at least one order of magnitude with an optimal parameter estimation.

6.3   Future work

The extension of these results is focused on two aspects, which are the optimisation and automation of the laser peak detector algorithm and the improvement in accuracy of the parameter estimation process in the calibration

algorithm. Both issues will lead to an increase in the 3D reconstruction accuracy.

6.3.1 Peak optimisation

The peak detection performance relies on the identification of a set of FIR filter parameters, which has been done *by hand* in our study. However, a significant contribution can be made toward the obtention of an automated procedure for the estimation of these filter parameters. Since the study of the frequency distribution of each image row signal has been carried out, it is clear that a good starting point is the FFT decomposition of these signals, analysing the worst case in terms of both cut-off frequency and transition band abruptness. In addition, once the above indicators have been identified, the filter parameters can be estimated using either a robust estimation method, like RANSAC, or a Maximum Likelyhood estimator such as the Fundamental Numerical Scheme or Levenberg-Marquardt, depending on the amount and nature of the noise present in the signals. The peak detection is the former step in the digitisation, hence significantly conditioning the subsequent processes. If this optimisation was undertaken an improvement in the rest of the shape digitisation process would become.

6.3.2 Improvement of parameter estimation

It has already been pointed that the better is the parameter estimation in the calibration procedure, the better and more reliable is the reconstruction. Total Least Squares (TLS) techniques has been prooved to be better than Algebraic Least Squares (ALS), however, the use of iterative methods are very useful for refining the results. The Fundamental Numerical Scheme (FNS), described in chapter 4, perfoms comparatively equal than TLS, although its authors [CBvdHG00] have prooved that it significantly improves the results of several parameter estimation problems such as conic or surface fitting or the Fundamental Matrix estimation, if covariance information is available for each data point involved. It is clear, then, that the calibration process should include a thorough study of covariance information for each application target, since covariance is a measure of how much and which type of noise is present in the image formation process. Covariance matrices cannot be induced by the camera manufacturer signal-to-noise specs, since it is strongly dependent on the type of surface and the lighting conditions, hence the great difficulty of estimating good covariances reliably.

6.4   Related publications

The results reported in this thesis were published in international conferences like IFAC, IROS or ICPR. In addition, the proposal of the new calibration method has been submitted to the Journal of Optical Engineering (a SPIE journal).

6.4.1   Journals

- Josep Forest, Joaquim Salvi, Enric Cabruja and Marc Bigas. Projective calibration of a laser 3D scanner using the complete quadrangle. submition acknowledged by the Journal of Computer Vision and Image Understanding in May 6th, 2004. In this article, the new calibration method has been described, together with the results and performance report.

- Marc Bigas, Enric Cabruja, Josep Forest and Joaquim Salvi. Review of CMOS Image Sensors. Submitted to the IEEE Transactions on Electron Devices. A survey on CMOS image sensors was published here. The scope was to get more in depth on CMOS image sensors to carry out an On-chip range computation system using this technology.

- Josep Forest, Joaquim Salvi, Enric Cabruja, Carles Pous. A flexible zero-crossing-based peak detector for laser stripe 3D scanners. Submitted to the Journal of Digital Signal Processing, Elsevier.

6.4.2   Conferences

- Josep Forest, Joaquim Salvi and Joan Batlle. Range Imaging System fo Underwater Applications. Proceedings of the 5th IFAC Manouvering Control for Marine Crafts, 2000. In this article, the test of the underwater range finder system was described.

- Josep Forest, Jordi PagÃ¨s, Joaquim Salvi and Josep Amat. 3D Image Ranging System for the Underwater Robot GARBÍ. Proceedings of the *IX Simposium Nacional de Reconocimiento de Formas y Análisis de Imágenes*, 2001. Novel results obtained with the underwater range equipment were published here.

- Josep Forest and Joaquim Salvi. A review of laser scanning three-dimensional digitisers. Proceedings of the Intelligent Robots and Systems 2002, pp. 73-78 The survey on laser scanners was published and presented.

- Josep Forest, Joaquim Salvi, Enric Cabruja and Carles Pous. Laser stripe peak detector for laser scanners: a FIR filter approach. to appear in Proceedings of the International Conference on Pattern Recognition, Cambridge 2004. The contribution of the new peak detector algorithm was published here.

- Josep Forest, Isidro RodrÃguez, Joaquim Salvi and Enric Cabruja. A Real-time low resolution range-finder for robotic applications. Proceedings of the Workshop on European Scientific and Industrial Collaboration, WESIC 2003. May, 2003. pp. 381-388. A report on the implementation of a low resolution discrete On-chip range computation was given. The results are collected in this thesis in appendix A

- Josep Forest, Josep Ma. Teixidor, Joaquim Salvi and Enric Cabruja. A proposal for laser scanners sub-pixel accuracy peak detector. Proceedings of the Workshop on European Scientific and Industrial Collaboration, WESIC 2003. May, 2003. pp 525-532. Together with the previous article, a former approach with early results on the performance and operation of the peak detector was published here.

Appendix A

Applications

In order to get deeper in how a three-dimensional digitiser works, some of the methods discussed in chapter 2 have been simulated using $Octave$[1]. The calibration procedures have been followed as faithfully as the level of abstraction allowed, and the performance has been evaluated making the three-dimensional reconstruction of a well known plane at four different ranges from the system (typically taking the image plane of the image sensor as a reference), and computing the maximum absolute reconstruction errors in the X,Y and Z directions.

A.1   The Chen & Kak system.

Figure A.1 shows the co-ordinate system arrangement which has been considered during the simulation. The relative pose between the camera and the laser emitter has been established such that the planes $X_L Y_L$ and $X_C Y_C$ are coincident, just for simplicity in the generation of the synthetic data. This does not affect in principle the performance of the system, nor the calibration procedure, since this method is transparent to explicit geometric considerations between those devices.

B is the separation or *base-line* between the centres of both the camera $\{C\}$ and the laser $\{L\}$ co-ordinate frames in the X direction. Note that the planes $X_L Z_L$ and $X_C Z_C$ are not coincident. **OLz** is the distance between the centres of the world $\{W\}$ and the laser $\{L\}$ co-ordinate frames. **R** is the range at which the calibration plane is located with respect to the world co-ordinate frame. **R** is known during the simulated calibration procedure, and takes the values from 0 to 110 mm in increments of 10 mm. The angle of projection (α) between the laser plane and Z_L is constant and set to 20 degrees.

---

[1]GNU Octave is a high-level language, primarily intended for numerical computations.
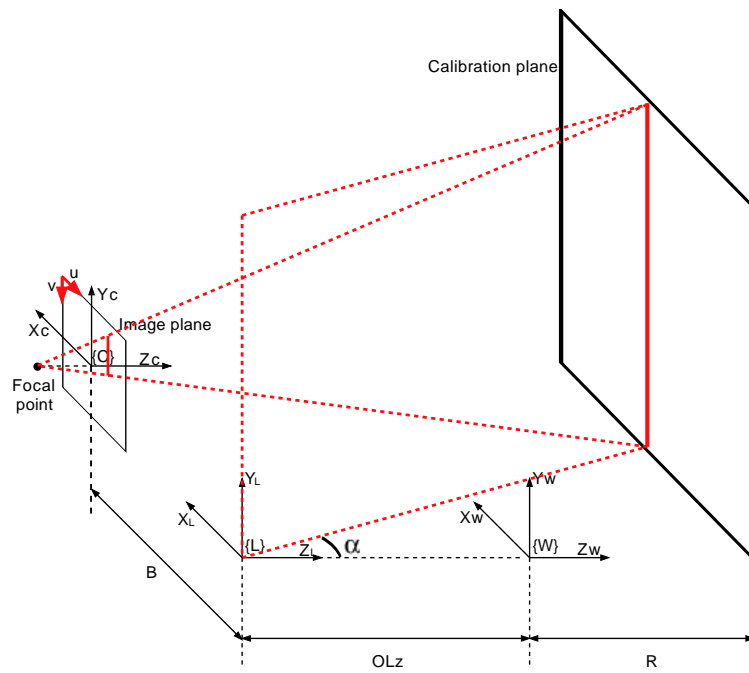
Figure A.1: Co-ordinate frames for Chen & Kak system simulation.
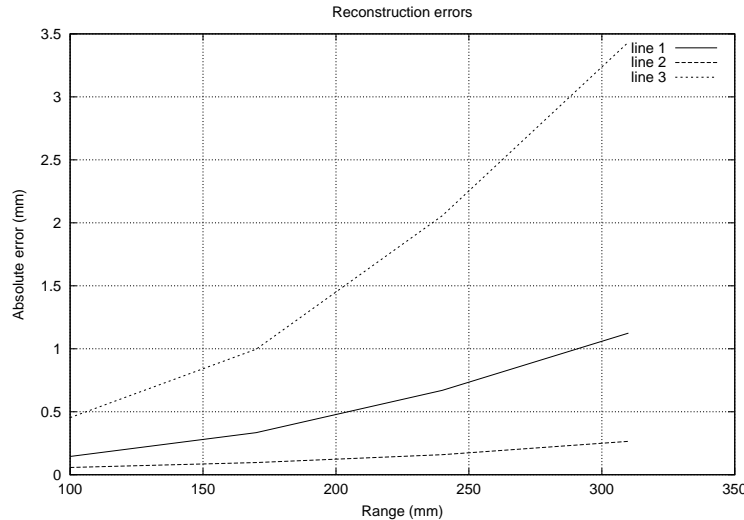
Figure A.2: Reconstruction errors as a function of range.

According to Hartley & Zisserman ([HZ03]), the performance evaluation of any three-dimensional system based on one or more cameras must take into consideration the existence of noise in the imaged data. This noise has been found to fit a Gaussian probability distribution.

In addition, Trucco et.al. ([TFFN98]) state that sub pixel accuracy for slit detection is strongly recommended for high accuracy three-dimensional reconstructions, hence it has been taken into consideration for simulation purposes. Table A.1 summarises the results obtained after the simulation of a calibration and reconstruction of a known plane. The four right-hand-side columns of this table correspond to the ranges at which the accuracy has been evaluated. These four distances at which the plane has been reconstructed are referred to the $\{C\}$ and $\{L\}$ co-ordinate frames, instead of $\{W\}$, in order to see more clearly what is the real distance of the object to the scanner. The added Gaussian noise has been chosen such that the maximum error in slit detection is of about 7 pixels, as shown in row number 2. Figure A.2 shows these results graphically. The upper curve (the largest error) corresponds to the error in range or Z direction, as it was expected, while the smallest error is found in the Y direction.

Figures A.3 and A.4 (a) to (d) show the reconstruction of a plane for Wz=0 and Wz=110 mm respectively. Parts (a) show a general view of the reconstructions, while parts (b) to (d) show different views of (a) where the errors in the Z, X and Y directions respectively can be appreciated in more detail. As can be seen, considering the error in pixels, the performance of the
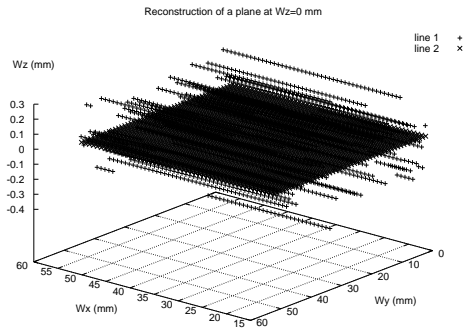
(a)
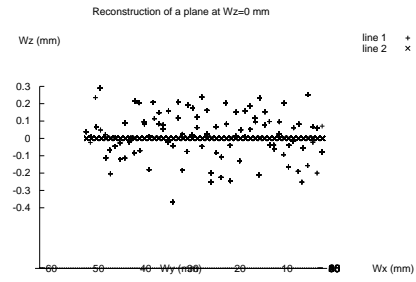
(b)
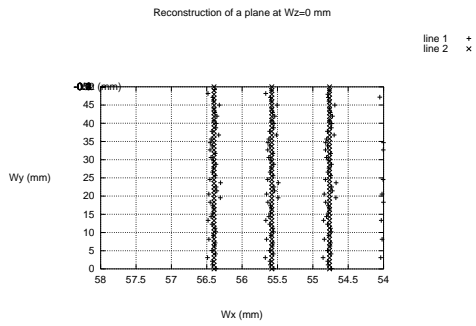
(c)
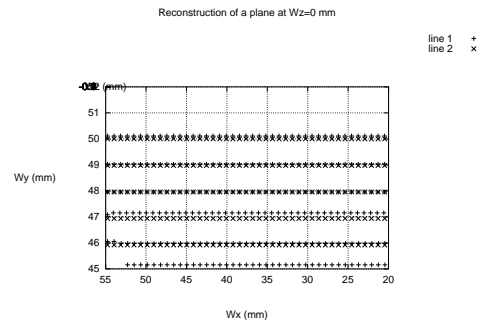
(d)

Figure A.3: Reconstruction of a plane at Wz=0.

Table A.1: Performance of the Chen and Kak scanner.

Range (mm)	100	170	240	310
Image error (pix.)	6,8916	7,0088	7,0505	7,0395
X error (mm)	0,144	0,3342	0,6703	1,1239
Y error (mm)	0,0573	0,0963	0,1594	0,2651
Z error (mm)	0,4535	0,9967	2,0569	3,4345
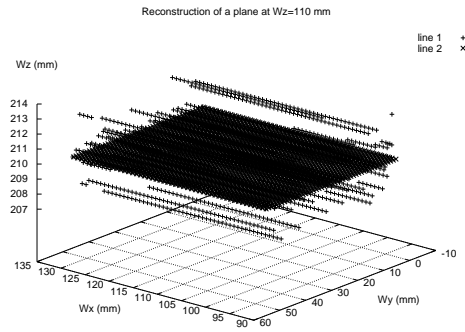
(a)
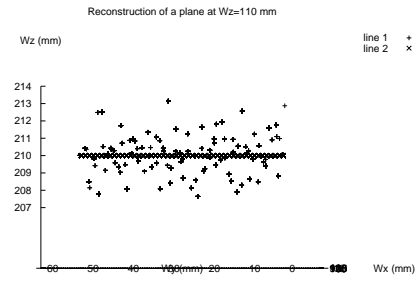
(b)
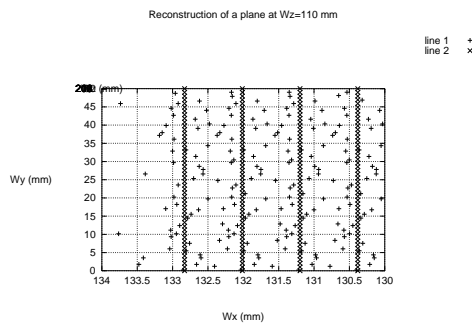
(c)
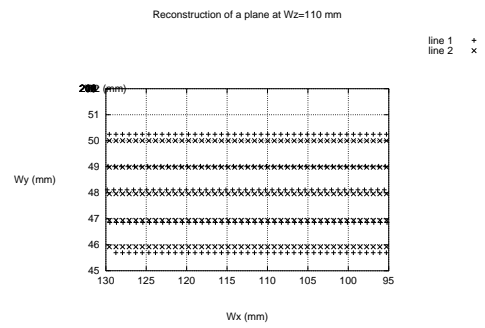
(d)

Figure A.4: Reconstruction of a plane at Wz=110.

system is considerably good, although it is expected to decrease when it is implemented using real devices. One of the reasons is that in this simulation, the stripe detection has been assumed to be carried out with a nearly ideal sub-pixel accuracy, that is, up to the machine accuracy, which is very high. In real systems based on discrete array cameras, the quantisation error is considerable, although sub-pixel accuracy peak detection techniques have shown to perform up to 1/8 - 1/9 of a pixel.

A.2 Smart sensors. Yet another approach

Figure A.5 shows the geometrical arrangement for the simulation of a hypothetical such device. A prototype based on the S7585 sensor of Hamamatsu has been built. This sensor is a simple 5x5 photo-diode array, hence the rest of the circuitry, both analog and digital, is assembled in separate PCBs. In addition, a detection scheme similar to that of Yokohama et. al.([YSYI94]) has been adopted and a new calibration method has been contributed, where just a simple known plane is required as a calibration target.

A.2.1 Slit detection

Yokohama et. al. used two photo-sensitive areas in each of the pixels in the array, one of them being slightly bigger in area than the other. The slit detection was achieved by comparison between the voltage outputs of both neighbour photo-diodes, hence, a detection signal was triggered at the instant in which the peak light intensity fell onto the middle of the two photo-diodes. Figure A.6 depicts the evolution of the two voltages with respect to time, showing the instant of detection. In our approach, the photo-diodes in the array have been used in couples, such that four detection "points" exist in each row. Hence the two voltages to be compared have been taken from the two neighbour pixels in each couple. Figures A.7 and A.8 show one row of the array, in the common cathode configuration (by manufacture default) and a simple comparator circuit which provides the detection signal, respectively. In addition, an offset voltage has been added in one of the comparator inputs in order to cancel out thermal noise, avoiding false detections.

A.2.2 Calibration procedure

According to figure A.5, the world co-ordinate frame is chosen to be placed in front of the {L} (laser) co-ordinate frame, with both Z_L and Z_W axis being coincident and the origins separated by OLz. The image sensor co-ordinate
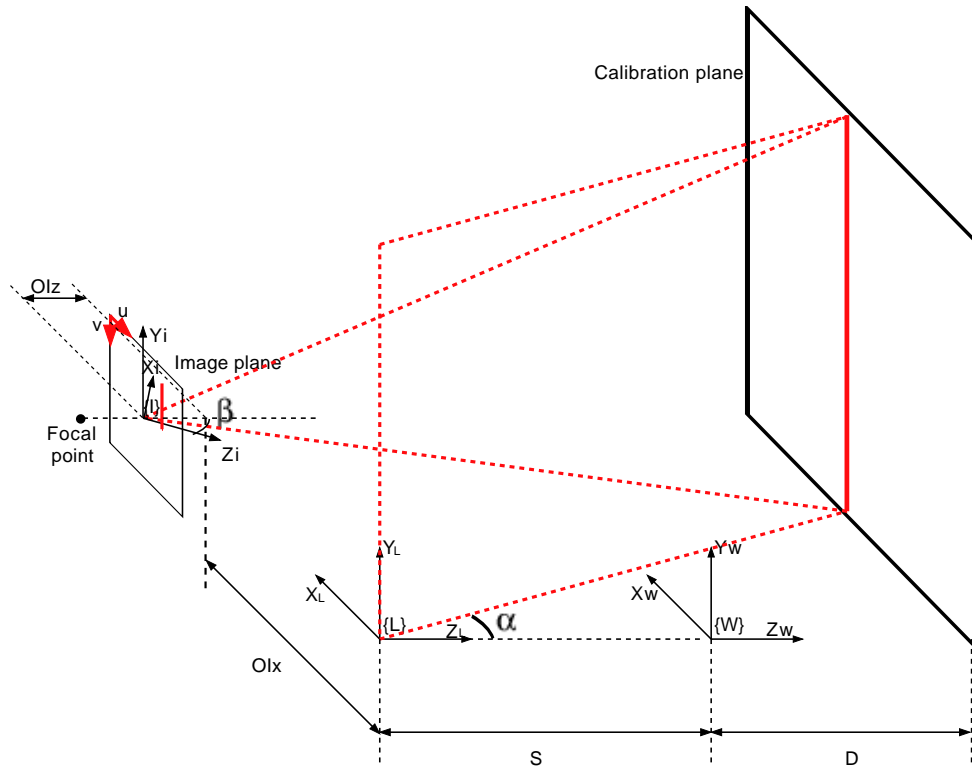
Figure A.5: Co-ordinate frames for Smart Sensor-based system simulation.
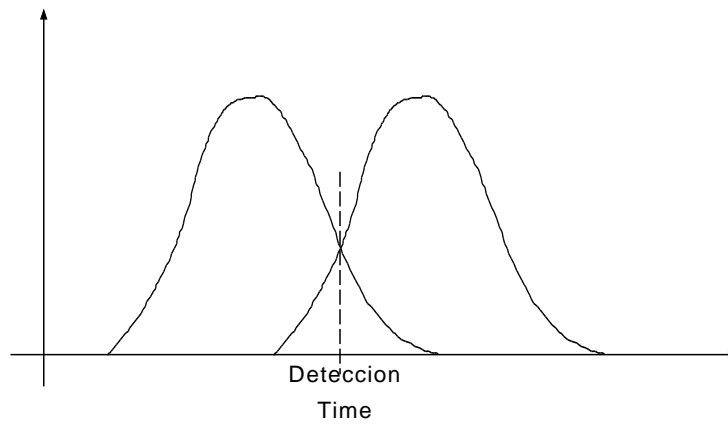
Figure A.6: Detection of laser slit by voltage comparison.
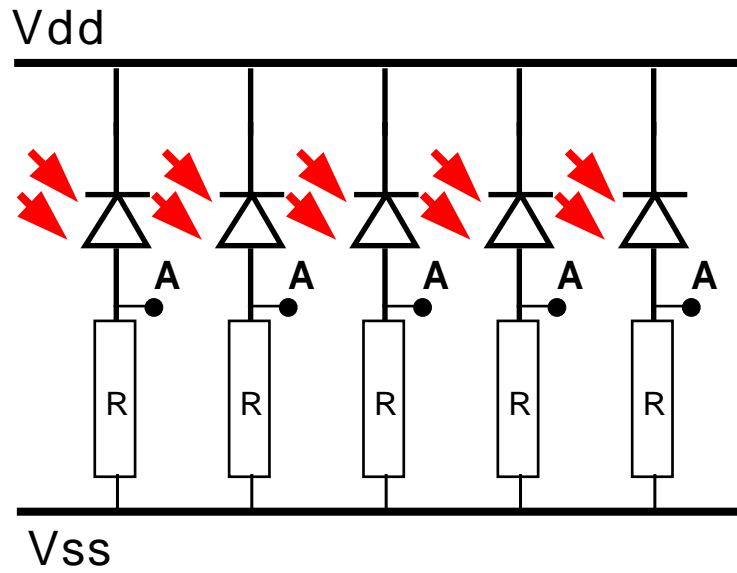
Figure A.7: One row in common cathode.
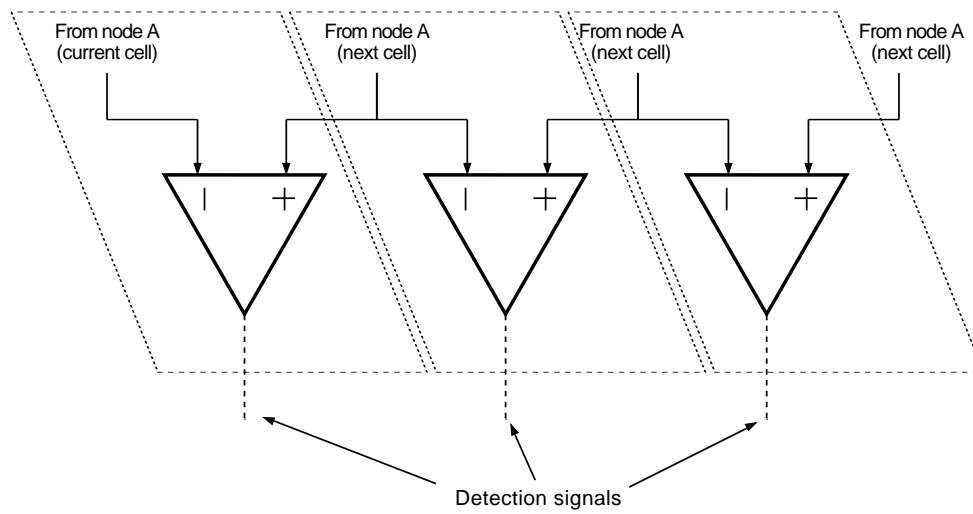
Figure A.8: Detection circuit.

frame {I} is considered to be separated from {L} by OIx and OIz along the X_L and Z_L axis respectively. Note that the $X_C Z_C$ and $X_L Z_L$ planes are supposed to be coincident. In addition, {I} has a non-zero orientation angle (β) with respect to the Z_W axe.

Assuming a pinhole model, a set of linear equations may be used in order to describe the behaviour of a camera, as stated and proved by [Fau93]. The pinhole model is valid if no zoom or wide angle lenses are used, since these lenses introduce radial as well as tangential distortion in the image, both of them non-linear effects. Since we are using a 50mm lens (a reflex camera body has been used as a housing for the 5x5 array), it is reasonable to assume the pinhole linear camera model for our purposes.

Camera model: intrinsic and extrinsic parameters

Any kind of discrete array-based video camera may be modelled in two stages, consisting in how the image points or pixels (in pixel co-ordinates) are related to an arbitrarily placed metric co-ordinate frame -*Intrinsic Parameters*-, and in how this metric co-ordinate frame is located with respect to an arbitrary world co-ordinate frame {W} -*Extrinsic Parameters*-.

The intrinsic parameters of a camera, assuming a pinhole model, are the focal distance (f), the principal point pixel co-ordinates (u_0, v_0) and the horizontal and vertical pixel pitch constants of proportionality (K_u, K_v). The reader is pointed to [Fau93] and [SAB02] to get deeper in camera calibration issues. In our case, though, we are not dealing with an already housed camera, but it is the camera what has been built, hence the horizontal and vertical constants of proportionality are supplied by the sensor manufacturer and the principal point must be well aligned with the optical axis of the lens, which is a mechanical issue, and has been chosen to be $(0, 0)$. Hence only f must be estimated in the calibration procedure, since we are working directly with the electric signals coming from each of the photo-diodes.

The estimation of the camera extrinsic parameters describe how the camera itself (or more concretely, the camera co-ordinate frame) is located with respect to the world co-ordinate frame. In our case, only the X (OIx) and Z (OIz) co-ordinates have been assumed to be variable (i.e. unknown a priori), since Y_C has been assumed to be parallel to both Y_L and Y_W. This assumption may be adopted if the laser plane emitter is perfectly vertical with respect to Y_L.

Calibration equations

The position and orientation of {I} with respect to {W} is described by the homogeneous transformation shown in equation A.1, hence, the point co-ordinates of the focal point, expressed as $^{I}FP = [0, 0, -f, 1]^{T}$ with respect to {I}, is obtained with respect to {W} making $^{W}FP =^{W} T_I \cdot^{I} P$, yielding the expression of equation A.2.

$$^{W}T_I = \begin{bmatrix} cos(\beta) & 0 & sin(\beta) & OIx \\ 0 & 1 & 0 & 0 \\ -sin(\beta) & 0 & cos(\beta) & OIz \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.1}$$

$$^{W}FP = \begin{bmatrix} OIx - f \cdot sin(\beta) \\ 0 \\ OIz - f \cdot cos(\beta) \\ 1 \end{bmatrix} \tag{A.2}$$

Once the focal point is expressed with respect to {W}, each of the detection point locations will be used in order to get the direction vectors (equation A.3) which, together with ^{W}FP, yield the expression of equation A.4, which is the line-of-sight equation in parametric form for each detection point, where R_X, R_Y, R_Z are the co-ordinates of the line-of-sight points, nx, ny are the pixel co-ordinates of each detection point, with respect to {I}, and dx, dy are the pixel pitch in the horizontal and vertical directions respectively. The pixel pitch is the inverse of the pixel pitch constant of proportionality (K_u, K_v). As can be seen in equation A.4, R_Y is dependent only on the Y co-ordinate in pixels (ny) and the pixel vertical pitch (py). This is due to the fact that there is no *vertical* information provided by the laser scan, since the laser shape is a vertical plane, which projects a slit parallel to the Y_W axis. Hence, the measurement error in the Y direction is ±7.5mm at a 500mm range, independently of the scanning mechanism, if the vertical detection point is assumed to fall in the centre of the pixels. Clearly, this shows that in order to minimise the reconstruction errors in the vertical, or Y, direction it is recommended that the pixels have the vertical pitch as short as possible.

$$V_d^W = \begin{bmatrix} nx \cdot dx \cdot cos(\beta) - f \cdot sin(\beta) \\ ny \cdot dy \\ nx \cdot dx \cdot sin(\beta) - f \cdot cos(\beta) \\ 0 \end{bmatrix} \tag{A.3}$$

$$
\begin{bmatrix} R_x \\ R_y \\ R_z \\ 1 \end{bmatrix} = \lambda \cdot \begin{bmatrix} nx \cdot dx \cdot cos(\beta) - f \cdot sin(\beta) \\ ny \cdot dy \\ nx \cdot dx \cdot sin(\beta) - f \cdot cos(\beta) \\ 0 \end{bmatrix} + \begin{bmatrix} OIx - f \cdot sin(\beta) \\ 0 \\ OIz - f \cdot cos(\beta) \\ 1 \end{bmatrix}
$$

(A.4)

In order to get the reconstruction (or estimated three-dimensional co-ordinates) of a point contained in a line-of-sight, it is necessary to obtain the point intersection of the laser plane with the line-of-sight itself. Hence, the laser plane equation must be obtained in terms of α, the scanning angle. $^{L}O = [0, 0, 0, 1]^{T}$ is the laser rotation centre point with respect to {L}. Since the orientation of {L} has chosen to be equal to the orientation of {W}, and its position is $-S$ mm away from the origin of {W} along Z_W, equation A.5 expresses $^{W}T_L$ by means of a homogeneous transformation. Hence, the laser rotation point (it is a rotation *axis*, indeed, but a *point* has been chosen for notation convenience only), can be found with respect to {W} by left-multiplyingg ^{L}O, such that $^{W}O =^{W} T_L \cdot^{L} O$, and its expression is shown in equation A.6.

$$
^{W}T_L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -S \\ 0 & 0 & 0 & 1 \end{bmatrix}
$$

(A.5)

$$
^{L}O = \begin{bmatrix} 0 \\ 0 \\ -S \\ 1 \end{bmatrix}
$$

(A.6)

The normal vector which describes the laser plane orientation may be obtained by left-multiplying the normal vector expression when $\alpha = 0$, by a rotation homogeneous transformation upon the Y_L axe, as it is expressed in equation A.7. In addition, if V_n^{L} is left-multiplied by $^{W}T_L$, the normal vector expression with respect to {W} is obtained, as shown in equation A.8. Since the orientation of {L} and {W} are equal, there have been no changes in the normal vector expression. Once both the normal vector and a point of the laser plane (which has been chosen to be the rotation centre for convenience) have been found, the laser plane equation in terms of α is configured, and its

expression is shown in equation A.9.

$$V_n^L = \begin{bmatrix} cos(\alpha) & 0 & sin(\alpha) & 0 \\ 0 & 1 & 0 & 0 \\ -sin(\alpha) & 0 & cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} cos(\alpha) \\ 0 \\ -sin(\alpha) \\ 0 \end{bmatrix} \tag{A.7}$$

$$V_n^W = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -S \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} cos(\alpha) \\ 0 \\ -sin(\alpha) \\ 0 \end{bmatrix} = \begin{bmatrix} cos(\alpha) \\ 0 \\ -sin(\alpha) \\ 0 \end{bmatrix} \tag{A.8}$$

$$X \cdot cos(\alpha) - Z \cdot sin(\alpha) = S \cdot sin(\alpha) \tag{A.9}$$

Once both the parametric equation of the line-of-sight and the laser plane equation have been found, a system of 4 linear equations can be set, but there is still a fifth equation which must be considered in order to obtain the calibration parameters. The calibration parameters are β, S, OIx, OIz and f, and they must be obtained from the knowledge of a calibration target object, which is, in our case, a plane located at known distances along the Z_W axis and parallel to the $X_W Y_W$ plane. Hence the fifth equation of the linear system is $Z = d_i$, where d_i is the distance of the i-th plane to the origin of {W}, as shown in equations A.10.

$$\begin{cases} (1)X \cdot cos(\alpha) - Z \cdot sin(\alpha) = S \cdot sin(\alpha) \\ (2)X = \lambda \cdot (nx \cdot dx \cdot cos(\beta) - f \cdot sin(\beta)) + OIx - f \cdot sin(\beta) \\ (3)Y = \lambda \cdot ny \cdot dy \\ (4)Z = \lambda \cdot (-nx \cdot dx \cdot sin(\beta) - f \cdot cos(\beta)) + OIz - f \cdot cos(\beta) \\ (5)Z = d_i \end{cases} \tag{A.10}$$

If λ is isolated from the A.10(2) and A.10(4) equations, two expressions for λ are obtained. If these two expressions are equated, the line-of-sight equation in the plane $X_W Z_W$ is obtained, as shown in equation A.11. Note that it is senseless to consider the line-of-sight equation in the vertical direction, since it does not contribute any additional information for obtaining the calibration parameters.

$$Z \cdot (nx \cdot dx \cdot cos(\beta) - f \cdot sin(\beta)) + X \cdot (nx \cdot dx \cdot sin(\beta) + f \cdot cos(\beta)) +$$
$$+OIz \cdot (f \cdot sin(\beta) - nx \cdot dx \cdot cos(\beta)) - OIx \cdot (f \cdot cos(\beta) + nx \cdot dx \cdot sin(\beta)) +$$
$$+f \cdot nx \cdot dx = 0$$

$$\tag{A.11}$$

If A.10(5) is substituted into A.10(1) and A.11, a new system of 2 linear equations is obtained, which, after isolating X in both of them and equating, equation A.12 is obtained, where the only unknowns are the calibration parameters.

$$OIz - OIx \cdot tan(\beta) + nx \cdot dx \cdot (OIz \cdot \frac{tan(\beta)}{f} + OIx \cdot \frac{1}{f} - sec(\beta)) -$$
$$- nx \cdot dx \cdot d_i \cdot \frac{1}{f} + d_i \cdot tan(\beta) - nx \cdot dx \cdot d_i \cdot tan(\alpha) \cdot \frac{tan(\beta)}{f} - \quad\quad (A.12)$$
$$- nx \cdot dx \cdot tan(\alpha) \cdot \frac{tan(\beta) \cdot S}{f} - tan(\alpha) \cdot S = d_i \cdot tan(\alpha)$$

Obtaining the calibration parameters

In order to obtain the calibration parameters (β, S, OIx, OIz and f), several complete scans on the calibration plane at different ranges must be made. In these scans, all the detection points **in a row** must detect the scanning stripe projected on the calibration plane. A complete scan will detect the scanning *times* in which the stripe *cuts* each of the line-of-sights, hence, in a given instant of time, when the stripe illuminates the **i-th** calibration plane, the laser light will intersect the **j-th** line-of-sight, as shown in figure A.9. Using i and j as the plane and line-of-sight indexes, equation A.12 can be arranged as shown in equation A.13.

$$\begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & dx \cdot n_j & -dx \cdot d_i \cdot n_j & d_i & -dx \cdot d_i \cdot n_j \cdot tan(\alpha_{ij}) & -dx \cdot n_j \cdot tan(\alpha_{ij}) & -tan(\alpha_{ij}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \cdot$$

$$\cdot \begin{bmatrix} OIz - OIx \cdot tan(\beta) \\ OIz \cdot \frac{tan(\beta)}{f} + OIx \cdot \frac{1}{f} - \frac{1}{cos(\beta)} \\ \frac{1}{f} \\ tan(\beta) \\ \frac{tan(\beta)}{f} \\ S \cdot \frac{tan(\beta)}{f} \\ S \end{bmatrix} = \begin{bmatrix} \vdots \\ d_i tan(\alpha_{ij}) \\ \vdots \end{bmatrix}$$

$$(A.13)$$

In order to solve the unknowns, a minimum of 7 *detections* or points should be provided for satisfying all the constraints, but since any real image capture system is noisy, according to [HZ03], it is much more advisable to include many more points so an overdetermined set of equations can be solved minimising some cost function. In our simulation, the calibration plane has been scanned at 4 different ranges, obtaining 4 points per scan, which yields a total of 16 points. In addition, since it has been proved sufficient above,
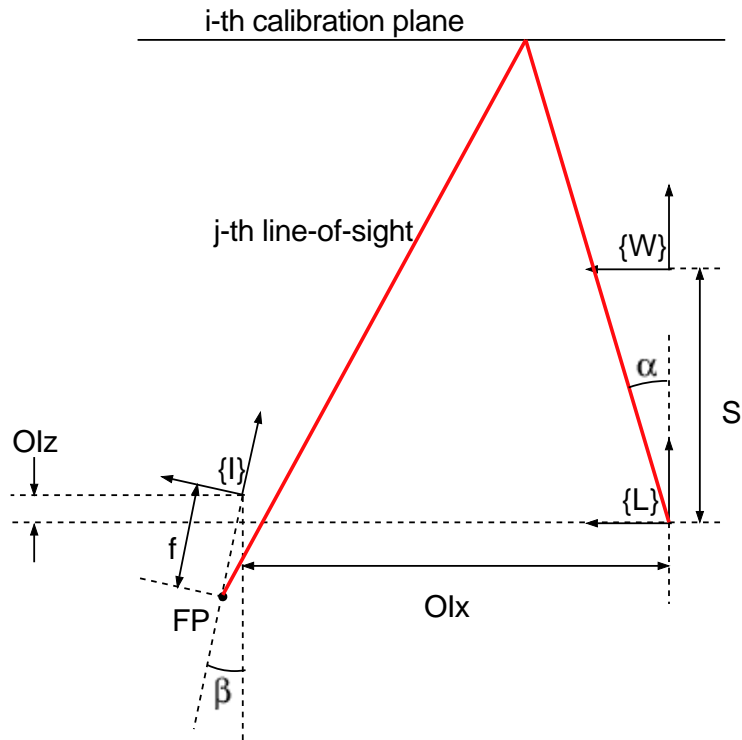
Figure A.9: Intersection between a line-of-sight and the laser plane.

only the row at $Y_I = 0$ has been used, and the singular value decomposition method has been utilised in order to solve the set of linear equations.

If the unknowns column vector is called U, with components $U_1 \ldots U_7$, the expressions of the calibration parameters can be obtained from A.13, as shown in equation A.14. The real parameters and the estimated ones do not show to be appreciably different, and their values in our simulation are: $\beta = 15^o$, $S = 200$ mm, $OIz = 20$ mm, $OIx = 400$ mm and $f = 50$ mm. Section A.2.3 reports a study of the noise influencing in the light detection and how it affects the reconstruction, while section A.2.4 shows the simulation results for different ranges and noise sources. The estimated noise power has been considered when the calibration parameters have been calculated.

$$
\begin{aligned}
\beta &= atan(U_4) \\
S &= U_7 \\
f &= \frac{1}{U_3} \\
OIz &= \frac{U_2 - U_1 U_3 U_4 + sec(\beta)}{U_3(1 + U_4^2)} \\
IOx &= U_1 + OIz \cdot U_4
\end{aligned}
\tag{A.14}
$$

A.2.3   Noise evaluation

Noise plays an important role in the results evaluation of any measuring device. Since laser scanners use to include at least 2 measuring device components (a camera and an angular sensor), it is necessary to evaluate which is the error introduced by both. In standard cameras, the most influencing noise has been found to follow a Gaussian probability distribution, which is a consequence of the point spread function, due to the imperfections in the lenses and the gray level digitisation. On the other hand, an angular sensor introduces electrical noise, as well as quantisation noise, either if this angular sensor is of a discrete nature or in the analog-to-digital stage conversion. In our case, the angular sensor is not used as a measure of the scanning angle itself, but it is the same image sensor, together with the additional circuitry, which provides a measure of the scanning angle. The angular sensor is used for constant speed control purposes only.

According to the above discussion, it is straightforward to spend a few paragraphs evaluating which kind of noise affects mainly our measuring device, and to which extent this noise distorts the measure. Since it is clear that it is the image sensor itself which provides the measure of the scanning angle, the noise evaluation study is restricted to the noise present in the image sensor, while the noise present in the angular sensor is assumed to be filtered out by the constant speed controller electronics. Under these assumptions,

Table A.2: The three main electrical noise sources under consideration.

Photon shot noise	$\#e^- = \sqrt{Q_{signal}/q}$
Dark current noise	$\#e^- = \sqrt{Q_{dark}/q}$
Johnson noise	$V_{RMS} = \sqrt{4 \cdot k \cdot T \cdot BW}$

two main noise sources have been found to influence the three-dimensional measure of our device: **electrical noise** and **quantisation noise**. The quantisation noise introduces a significant error in the measure, due to the truncation which the number of bits employed in the angular measure digitisation induces. It has been proved that this error is more important for low bit numbers, but when the number of bits rises above a particular threshold, the electrical noise becomes more important in the measure distortion.

Electrical noise, which has a uniform probability distribution, has shown to be of a great importance in image formation quality. Dierickx [Diea] and [Dieb] described the different noise sources which influence in image sensors in general but in CMOS-based image sensors particularly. Since we have used a CMOS photo-diode array, it has been considered of a great convenience to follow Dierickx's performance evaluation given in the CMOS image sensors-concepts and limits SPIE Photonics West 2000 Short Courses. The influencing electrical noise, in our sensor, can be decomposed into 3 categories: **Photon shot noise**, **Dark current noise** and **Thermal or Johnson noise**. Photon shot noise is due to the statistical variation on the number of absorbed/converted/collected photons (converted to electrons or charge carriers), it is considered as a fundamental noise source and it is device-independent, i.e there is no way to cancel it out, but it is inherent to the photon to charge conversion. Dark current is the apparent current present in the photo-diode under darkness conditions. The presence of dark current introduces the so called dark current noise, which is the statistical variation on the number of generated/recombined/collected charge carriers. Finally, the thermal or Johnson noise is the statistical variation of charge carriers agitation due to the temperature. Table A.2 collects the algebraic expressions of these 3 noise sources. In order to have a thorough performance evaluation, the total signal-to-noise ratio of the sensor under normal operation must be taken into account. This total noise is the sum of the noise induced by the three noise sources independently, since the principle of superposition applies. The device specifications as well as the signal-to-noise ratio evaluation are shown in table A.3, where the laser power has been cho-

sen to be of 6 mW peak, and only the case in which the laser is projected on a dark object (20% of reflectivity).

Table A.3: Performance evaluation in terms of the S/N.

Light features		
Distance of performance evaluation	0.3	[m]
Aperture angle	85	[deg]
Laser emitter peak light power	6	[mW]
Major axis	549.798704410454	[mm]
Minor axis	1.5	[mm]
Slit area	0.0025908653560936	[m2]
Light flux (direct)	2.31582856511175	[W/m2]
Light flux (reflected from white object)	1.8526628520894	[W/m2]
Light flux (reflected from dark object)	0.463165713022351	[W/m2]
Sensor features		
Pixel size	1300	[um]
Pixel area (square pixel)	0.00000169	[m2]
Pixel power	0.00000782750055007773	[W]
Photo sensitivity	0.39	[A/W]
Pixel current	0.00000305272521453031	[A]
Integration time	0.000065	[s]
Photo charge	0.000000000019842713894447	[C]
Photo charge	124016961.840294	[e-]
Pixel capacitance	0.00000000001	[F]
Pixel voltage	1.9842713894447	[V]
Noise Calculation		
Photon shot noise	11136.2903087291	[e-]
Dark currrent	0.05	[nA]
Dark current noise	142.521928137392	[e-]
Signal-to-noise ratio		
Total noise (shot+dark current)	11278.8122368665	[e-]
S/N (Photo charge/noise electrons)	10995.5693237738	Linear
S/N (Photo charge/noise electrons)	80.82435441977	(dB)
S/N considering thermal noise		
Noise equivalent voltage (shot+dark)	0.000180460995789864	[V]
Thermal (Johnson) noise	0.00000122758299108451	[V]

Table A.4: Metric errors due to electric noise.

Zw (mm)	10	110	210	310
dX (um)	+4,37E-01	+1,08E+01	+2,93E+01	+5,53E+01
dZ (um)	+4,93E-01	+1,21E+01	+3,31E+01	+6,24E+01

Total noise voltage	0.000180583754088973	
S/N (Pixel voltage/Noise total voltage)	**10988.0946902182**	Linear
S/N (Pixel voltage/Noise total voltage)	**80.8184478661937**	(dB)

The laser stripe has been modelled as an elongated ellipse, according to the manufacturer's data. The sensor features have been picked from the photo-diode array data-sheet.The photo charge is the generated charge (or current) in the photo-diode, during the exposition (or integration) time, which is the time during which a laser stripe image illuminates one pixel. As can be observed, the S/N does not change considerably either accounting for thermal noise or not, hence, a signal-to-noise ratio of **80.82** dB has been considered in our simulations.

A.2.4   Simulation results

Table A.4 show the results of the reconstruction of a plane for different ranges, with no errors due to quantisation. Table A.5 show the reconstruction errors due only to quantisation, for different quantisation bits. As can be observed, the error due to quantisation becomes smaller and smaller as the number of quantisation bits increase, but when using 16 bits or more, errors due to electrical noise become more important, hence, it does not make sense to include more than 24 bits, for the quantisation, since, as can be seen in table A.6, the errors stabilise for 16, 24 and 32 bits.

A.3   An attempt to Underwater Range Sensing

A.3.1   Introduction

In the last few years, the development of AUVs has experienced a great up-surge. There is no doubt that there's a lot of interest in developing robots

Table A.5: Metric errors due to quantisation.

	Zw (mm)	10	110	210	310
8 bits	*dX (um)*	+1,73E+03	+1,19E+03	+2,47E+03	+3,74E+03
	dZ (um)	+1,88E+03	+1,95E+03	+3,05E+03	+4,04E+03
16 bits	*dX (um)*	+2,70E+00	+1,28E+01	+2,59E+00	+2,80E+01
	dZ (um)	+5,93E+00	+1,45E+01	+5,69E+00	+3,16E+01
24 bits	*dX (um)*	+1,14E-02	+3,87E-02	+6,29E-02	+4,85E-02
	dZ (um)	+1,98E-02	+4,37E-02	+7,10E-02	+5,47E-02
32 bits	*dX (um)*	+9,44E-05	+1,81E-04	+2,88E-04	+4,32E-04
	dZ (um)	+1,06E-04	+2,04E-04	+3,25E-04	+4,87E-04

Table A.6: Metric errors due to both quantisation and electric noise.

	Zw (mm)	10	110	210	310
8 bits	*dX (um)*	+1,73E+03	+1,19E+03	+2,47E+03	+3,73E+03
	dZ (um)	+1,88E+03	+1,94E+03	+3,05E+03	+4,03E+03
16 bits	*dX (um)*	+8,53E+00	+6,96E-01	+1,42E+01	+4,76E+01
	dZ (um)	+9,62E+00	+2,91E+00	+1,60E+01	+5,36E+01
24 bits	*dX (um)*	+4,20E-01	+1,03E+01	+2,82E+01	+5,29E+01
	dZ (um)	+4,74E-01	+1,16E+01	+3,17E+01	+5,97E+01
32 bits	*dX (um)*	+4,37E-01	+1,08E+01	+2,93E+01	+5,53E+01
	dZ (um)	+4,93E-01	+1,21E+01	+3,31E+01	+6,24E+01

to carry out the unpleasant and dangerous underwater tasks which, nowadays, are done by humans. With this idea in mind, most of the efforts are focused on visual servoing, that is, on self-positioning the robot with respect to an object by using computer vision. It is well-known that underwater visual servoing becomes more difficult in unstable sea movement. The task of grabbing proper images under such circumstances is the primary problem which has to be solved in order to carry out complicated tasks such as object manipulation. Some research has been done with the aim of doting the Vortex underwater vehicle by a visual sensor based on the projection of a laser beam slit [CMRt96]. This sensor has been used in underwater pipe inspection [RB97]. However, the main problem of single slit projection is the reduced area which can be analysed. Of course, this problem can be solved by acquiring more scene information by the projection of more than a single slit, forming a stripped or grid pattern [Jar83]. However, the correspondence problem between captured and projected slits has then to be taken into account [RSB98]. Our approach is based on using a scanning laser slit. Using this approach the problem of correspondence is overcome by considering one slit projection at a time among a set of projections over time. In this way, time and the fact that there can't be more than one line on the same row of the image identifies each pixel uniquely.

The use of acoustic active systems has not been considered due to their slow acquisition rates and lack of accuracy compared to the thickness of a lase line. Acoustic imaging is very useful when building surface maps of the bottom because a simple cheap device can be mounted under a ship's keel. Map building is achieved by ordering the ship a straight rhumb.

Acoustic systems are not suitable for reconstruction of moving objects. Furthermore, static bottom-locked acoustic devices are very useful for position detection of *any thing* within a determined area, but they are useless if we are intending to build a system that can be carried by the robot itself in order to be positioned with respect to any natural element such as rocks or wall prominences or even man made elements such as the mast of an immersed ship.

The presented sensor will be mounted on our ROV called GARBI. GARBI is a project financed by the Spanish government and jointly developed between the Polytechnical University of Catalonia and the University of Girona. The robot has been primarily used in tele-operated tasks like underwater inspection [ACCP95]. It has also been used in underwater object manipulation by using its two tele-manipulated arms.

This paper is structured as follows. Firstly, the underwater robot is briefly introduced. Then, the camera and laser beam modelling and calibrating is explained in section 3. Section 4 includes the principle used to reconstruct

Figure A.10: The underwater robot GARBÍ.

the scene. Section 5 shows the experimental results. The paper ends with conclusions and further work.

A.3.2 The Underwater Robot

The vehicle developed was conceived for exploration in waters up to 200 meters. The robot has 4 degrees of freedom, corresponding to the linear movements of advancing, descending, turning, and pitching. A picture of this robot, while performing a test, can be seen in figure A.10. These movements are achieved by 4 propellers. With the aim of building a low cost underwater vehicle, GARBI was constructed using low cost materials, such as fibre-glass and epoxy resins. To solve the problem of resistance to underwater pressure, the vehicle is servo-pressurised to the external pressure using a compressed air bottle like those used in SCUBA diving. Air consumption is required only in vertical displacements during which the decompression valves release the required amount of air to maintain the internal pressure at the same level as the external one. The vehicle is equipped with two arms, thus it can, through tele-operation, perform some tasks of object manipulation [ACCP95]. The robot interior has been organised so as to install the following equipment: (1) Laser beam projecting system, (2) RGB Camera, (3) PC computer, (4) Power interface for the thrusters and arm motors and (5) Batteries.

A.3.3   The use of light underwater: Scattering and Absorption

The use of light in underwater environments is subject to the consideration of physical phenomena such as scattering and absorption, both of them present in the surface too, but considered only when very long distances or dusty environments are involved for measurement purposes in free air.

An exhaustive thorough study of light scattering by particles can be found in [VdH81].

Scattering becomes a major factor in underwater illumination systems due to the large amount of particles present within the water. This phenomena causes a change in light *Intensity*, *Polarisation* and *Phase*, although the loss of intensity is the most relevant drawback to be taken into account, due to the limitation it imposes to the distance the light can achieve. Scattering is the fact that light is reflected on the multiple floating particles within the water. Scattering can be observed too in free air when light passes through the mist or a cloud. Figure shows scattered light due to a particle is characterised by the scattering angle θ.

Absorption of light is a relevant aspect as well. It is much more important underwater than in free air for the same reason than scattering. Absorption takes place when part of the incident light is *absorbed* by a particle. This absorbed light, that can be thought of as energy, is transformed to some other form of energy (heat, for example), so the remaining light has less intensity than the incident one.

The combination of both scattering and absorption phenomena makes the light suffer an Extinction, which is the total absence of a particular wavelength at a determined distance from the focus of light. As it is stated in [VdH81]*:*
Extinction=scattering+absorption.

In order to avoid an early extinction of the light, more powerful light sources should be used, as well as choosing a coherent light source (laser) centred at a proper wavelength. Red is rapidly extinguished with distance underwater, but blue is the later wavelength to disappear. In our system, a green ($\lambda = 510nm$, 5mW output power) semiconductor laser has been chosen as a compromise between extinction and cost effectiveness.

A.3.4   The Range Imaging System

The laser beam rotates by means of a DC motor in order to scan the whole camera scope getting a reconstruction of the front scene. A green laser beam was chosen because it is less affected by scattering than a red one, thus

enlarging the camera scope and facilitating accurate segmentation results under no light control scenarios. The power control card was adapted to the new drivers technology and integrates the supply, the motor PWM control and the incremental encoder used to obtain the angular position of the laser beam. Furthermore, a processing card was used in order to segment the image and obtain the 2D co-ordinates of the laser beam in the image plane at video rate. The processing card converts the space colour from RGB to HLS and segments the laser light from the rest of the scene by using LUT memories [Bat93]. Thereafter, the card processes the video lines with the aim of obtaining the 2D co-ordinates. The first co-ordinate is given by the video line explored by the card, the second co-ordinate is given by computing the gravity centre of the segmented pixels. The card also synchronises the camera with the laser and gets the angular position of it for each image shot. Once the whole image has been processed the information is sent to the computer which produces the 3D reconstruction.

Camera Modelling and Calibration

The pinhole model is based on reducing the complexity of camera behaviour to an image plane Π where the scene is projected through the focal point O_c [Fau93]. The focal point is then placed at a fixed distance f from the image plane eliminating lens distortion. Then, giving a 3D object point P its 2D image projection p is obtained by the intersection of the optical ray v defined by P and O_c and the image plane Π. This model is defined by two transformation matrices ${}^C T_W$ and ${}^I T_C$. ${}^C T_W$ deals with the extrinsic parameters which express the position and orientation of the world co-ordinate system $\{W\}$ with respect to the camera co-ordinate system $\{C\}$. ${}^I T_C$ models the intrinsic parameters which express the optical characteristics and internal geometry of the camera. That is, the projection of the focal point O_c on the image plane (u_0, v_0) and the (α_u, α_v) parameters which permit the transformation from the camera co-ordinate system in millimetres to the image co-ordinate system in pixels.

The camera model determines the relationship between the 3D object point P and its 2D image point projection p by using the equation (A.15).

$$\begin{bmatrix} s^I X_i \\ s^I Y_i \\ s \end{bmatrix} = {}^I T_C \, {}^C T_W \begin{bmatrix} {}^W X_i \\ {}^W Y_i \\ {}^W Z_i \\ 1 \end{bmatrix} \tag{A.15}$$

where ${}^C T_W = \begin{bmatrix} R_{3x3} & t_{3x1} \\ 0_{1x3} & 1 \end{bmatrix}$

and $^{I}T_{C} = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$

Camera calibration is based on obtaining the intrinsic and extrinsic parameters of the camera model from a set of 3D object points with respect to the world co-ordinate system and their 2D image points with respect to the image co-ordinate system. The n 3D object points are obtained by using a calibrating pattern in which the 3D components can be easily measured. The 2D image points are obtained by measuring the projection of the 3D object points on the image plane of the camera. By using the equation (A.15) and arranging the unknowns the equation (A.16) is obtained, in which B is a vector of $2n$ components which contain the 2D image point components. Q is a matrix of $2n$ rows and 10 columns expressing the relationship between 2D and 3D points; and X is the vector of the 10 camera model unknowns which can be obtained by using the equation (A.17) [Sal97].

$$B = QX \tag{A.16}$$

$$X = \left(Q^T Q\right)^{-1} Q^T B \tag{A.17}$$

The more points we use to calibrate the camera the more accuracy is obtained. However, we also have to take care with the accurate measurement of the 3D object points of the calibrating pattern and their 2D image point projections.

Laser Model and Calibration

In order to get accurate reconstructions of the scene, the laser geometry (shape, position and orientation) with respect to a reference coordinate system $\{W\}$ must be known. To this aim, the system has been modelled as shown in figure A.11. In this model no constraints in the position, orientation or movement of the laser with respect to $\{W\}$ have been considered. The laser slit has been considered as a plane which rotates around the Z_L axis of the laser coordinate system $\{L\}$.

Within these considerations, up to 8 parameters have to be considered:

- X, Y and Z determine the position of the laser coordinate system $\{L\}$ with respect to the reference coordinate system $\{W\}$.

- α is the angle of rotation around Z_L, measured from X_L, while δ is the offset angle of rotation of $\{L\}$ with respect to Z_L, due to the mechanical misalignment between X_L and the mechanical origin of the encoder.
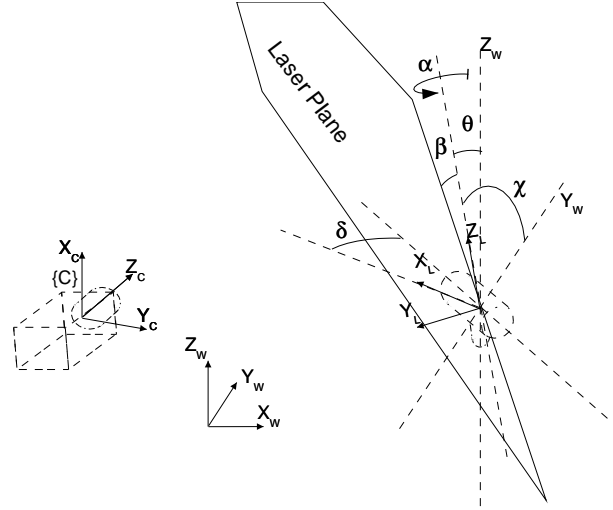
Figure A.11: Laser system with the 8 parameters describing its model.

- β is the angle that the plane forms with Z_L, measured around X_L, due to the misalignment of the laser emitter with respect to the motor's axis of rotation.

- θ and γ are the pan and tilt angles of rotation of $\{L\}$ with respect to X_W and Y_W respectively considering $\{L\}$ and $\{W\}$ coincide, due to errors in screwing the motor on the platform.

Among these 8 parameters, only α is known in each instant of time, because it is delivered by the encoder, and is the only one that doesn't have to be found in the calibration process.

A plane in R^3 is defined by a normal vector $(\overrightarrow{v}_n)$ and a point contained on the plane. In order to get $\overrightarrow{v}_n$ it is first necessary to compute ${}^W T_L$ which expresses the orientation of $\{L\}$ with respect to $\{W\}$, and is achieved by considering a rotation of δ around Z_W, followed by a rotation of θ around X_W, and then a rotation of γ around Y_W.

$\overrightarrow{v}_{nL}$ (normal vector to the plane with respect to $\{L\}$) is computed by a rotation of β around X_L, followed by a rotation of α around Z_L, considering it is normal to the plane $Y_L = 0$ when $\alpha = \beta = 0$.

In order to get $\overrightarrow{v}_{nW}$ (the normal vector with respect to $\{W\}$), it is necessary to compute $\overrightarrow{v}_{nW} = {}^W T_L \cdot \overrightarrow{v}_{nL}$.

Now, if we consider the restrictions $\theta = \gamma = 0^o$, for the camera and the laser system are rigidly joined, we get the final expression for $\overrightarrow{v}_{nW}$ in equation (A.18).

$$\vec{v}_{nW} = \begin{bmatrix} -\sin\alpha\cos\beta\cos\delta - \cos\alpha\cos\beta\sin\delta \\ -\sin\alpha\cos\beta\sin\delta + \cos\alpha\cos\beta\cos\delta \\ -\sin\beta \\ 0 \end{bmatrix} \tag{A.18}$$

The laser calibration is achieved by projecting the laser slit on the calibration pattern used for calibrating the camera. Segmenting the laser colour (green) at video rate through a special purpose video pre-processing card [Bat93], we get the line projection of the laser plane. After thinning the line, we have the 3D points corresponding to the laser line points. Now these points can be used in order to calibrate the laser.

As the 3D points belong to the laser plane, they can be used to get the laser plane equation $Ax + By + Cz + D = 0$.

Where A,B and C are the three components of the normal vector $\vec{v}_{nW}$, and D is defined as the dot product of a point P in the plane and $\vec{v}_{nW}$.

The plane equation can be operated and arranged in matrix form, obtaining equation (A.19).

$$\begin{bmatrix} & \vdots & \vdots & \\ -\sin\alpha_i & \cos\alpha_i & P_{xi}\cos\alpha_i + P_{yi}\sin\alpha_i & 1 \\ & \vdots & \vdots & \end{bmatrix} \begin{bmatrix} X + Y\tan\delta \\ Y - X\tan\delta \\ \tan\delta \\ \frac{\tan\beta}{\cos\delta}(P_z - Z) \end{bmatrix} = \begin{bmatrix} \vdots \\ P_{yi}\cos\alpha_i - P_{xi}\sin\alpha_i \\ \vdots \end{bmatrix}$$
$$\tag{A.19}$$

As the laser can sweep over the whole calibration pattern, multiple sets of points can be obtained, so sweeping from $\alpha = 0$ to $\alpha = 41^o$, we have one set of points corresponding to $P_z = 0$ and another one for $P_z = 100$. The system can be solved by considering two sets of only 4 points, but if more points are obtained, the error due to noise in the CCD of the camera, the segmentation process or the thinning algorithm can be minimised. Using the pseudo-inverse method, the solution vector $S = \begin{bmatrix} X + Y\tan\delta & Y - X\tan\delta & \tan\delta & \frac{\tan\beta}{\cos\delta}(P_z - Z) \end{bmatrix}^T = \begin{bmatrix} S_1 & S_2 & S_3 & S_4 \end{bmatrix}$ is obtained for each set of points. Using only one set of points, X, Y and δ can be obtained with the first 3 components of S, and using the results of S_4 for the two sets, we can find a solution for β and Z. The expressions for X, Y,δ,$\beta$and Z are shown in expressions (A.20) to (A.23).

$$X = \frac{S_1 - S_2 S_3}{1 + S_3^2} \tag{A.20}$$

$$Y = S_2 + X S_3 \tag{A.21}$$

$$\delta = \arctan(S_3) \tag{A.22}$$

$$\beta = \arctan\left(\frac{S1_4 - S2_4}{(P_{x0} + P_{x100})\cos\delta}\right)$$

$$Z = \frac{S1_4 \cos\delta}{\tan\beta} - P_{x0} \tag{A.23}$$

Finally, the 7 calibration parameters for the laser structure have been obtained. The procedure for obtaining them is a linear algorithm, so it is fast and more reliable as the number of 3D points increases.

A.3.5 Experimental Results

An experimental platform was built, incorporating the hardware system: camera, laser and coupling mechanisms. The camera and laser was modelled and calibrated by using two orthogonal calibrating planes with 20 equally spaced black squares on each one. The vertexes of these squares were accurately measured and used as the 3D object point set for the camera calibration. The laser system was calibrated by projecting the laser slit on both calibrating planes and measuring its 3D position, as the camera had already been calibrated. When both systems are calibrated, the system can be used for 3D reconstruction. Simple experiments were undertaken underwater, using a parallelepiped. The length of the slanting side of the parallelepiped is 82mm, measured with a 10μm accuracy calliper, while the reconstruction yield a result of 79.6mm, so an error of 3% has been committed.

Appendix B

Notation

This appendix synthesizes the nomenclature used to express coordinate systems and transformation matrices within the thesis.

B.1 Mathematics Convention

$\{H\}$ defines a coordinate system H, which is composed of an origin O_H and either two $\{X_H, Y_H\}$ or three $\{X_H, Y_H, Z_H\}$ orthonormal axis, depending on the number of dimensions defined. Hence, a 3D coordinate system $\{H\}$ can be expressed as $\{H\} = \{O_H, X_W, Y_W, Z_W\}$, and the same applies for a 2D coordinate system, removing the Z_W axis.

Each point P is always expressed as a column vector and is related to a coordinate system. Hence, $^H P$ relates the point P with respect to $\{H\}$, where $^H P = ^H [X, Y, Z]^T$.

In addition, extending to projective geometry, the following notation for homogeneous coordinates applies:

- $^H P = ^H [X, Y, Z, W]^T$ expresses a point in $\mathbb{P}^3$

- $^H P = ^H [x, y, z, 1]^T = [X/W, Y/W, Z/W, 1]^T$ expresses a point $^H P \in \mathbb{R}^3$, in homogeneous coordinates.

- $^H P = ^H [X, Y, W]^T$ expresses a point in $\mathbb{P}^2$.

- $^H P = ^H [x, y, 1]^T = [X/W, Y/W, 1]^T$ expresses a point $^H P \in \mathbb{R}^2$, in homogeneous coordinates.

- $^H V = ^H [X, Y, Z, 0]^T$ is a vector of $\mathbb{R}^3$.

- $^H V = ^H [X, Y, 0]^T$ is a vector of $\mathbb{R}^2$.

Using this notation, it is easy to express rigid transformations of coordinate systems:

- $^{J}K_{H}$ expresses the position and orientation of $\{H\}$ with respect to $\{J\}$.

- The inverse of a rigid transformation is expressed by exchanging super- and sub-indices, i.e. $(^{J}K_{H})^{-1} = {}^{H}K_{J}$.

- The composition of a series of rigid transformations is expressed with the superindex of the former transformation and the subindex of the last one, i.e. $^{A}K_{B} \cdot {}^{B}K_{C} \cdot {}^{C}K_{D} = {}^{A}K_{D}$.

Finally, a rigid transformation $^{J}K_{H}$ is a 4×4 homogeneous matrix such as:

$$^{J}K_{H} = \begin{pmatrix} ^{J}R_{H3x3} & ^{J}T_{H3x1} \\ 0_{1x3} & 0_{1x1} \end{pmatrix}$$

where $^{J}R_{H}$ expresses the orientation of the axis of $\{H\}$ with respect to $\{J\}$, and $^{J}T_{H}$ is the translation of O_{H} with respect to $\{J\}$.

Bibliography

[AB73]   Gerald J. Agin and Thomas O. Binford. Computer description of curved objects. In *International Joint Conferences on Artificial Intelligence*, pages 629–640, 1973.

[ACCP95]   J. Amat, J. Codina, X. Cufí, and J. Puigmal. Vision based control of underwater vehicle for long duration observations. In *Int. Conf. on Autonomous Robots*, volume 1, pages 273–277, 1995.

[AN87]   B.F. Alexander and K.C. Ng. 3d shape measurement by active triangulation using an array of coded light stripes. In *Proceedings of the SPIE: Optics, Illumination and Image Sensing for Machine Vision II*, volume 850, pages 199–209, 1987.

[AS03]   Xavier Armangué and Joaquim Salvi. Overall view regarding fundamental matrix estimation. *Image and Vision Computing*, 21(2):205–220, 2003.

[ASP87]   Kazuo Araki, Yukio Sato, and Srinivasan Parthasarathy. High speed rangefinder. In SPIE, editor, *Optics, illumination and image sensing*, volume 850, pages 184–188, 1987.

[Bac00]   Benjamin T. Backus. Stereoscopic vision: what's the first step ? *Current Biology*, 10:R701–R703, 2000.

[Bat93]   J. Batlle. *Aportació a la detecció de moviment independentment del moviment de la càmera*. PhD thesis, Universitat Politècnica de Catalunya, 1993.

[BBvEC99] Benjamin T. Backus, Martin S. Banks, Raymond van Ee, and James A. Crowell. Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Research*, 39:1143–1170, 1999.

[BCGdH00]   Michael J. Brooks, Wojciech Chojnacki, Darren Gawley, and Anton Van den Hengel. Is covariance information useful in estimating vision parameters? In Sabry F. El-Hakim and Armin Gruen, editors, *Proceedings of SPIE. Videometrics and Optical Methods for 3D shape measurement*, pages 195–203, December 2000.

[BCGvdH01] M.J. Brooks, W. Chojnacki, D. Gawley, and A van den Hengel. What value covariance information in estimating vision parameters? In *Proceedings of the Eighth Computer Vision, ICCV 2001*, volume 1, pages 302 – 308, July 2001.

[Bes88]   P.J. Besl. Active, optical range imaging sensors. *Machine Vision and Applications*, 1:127–152, 1988.

[BFPH01]   Benjamin T. Backus, David J. Fleet, Andrew J. Parker, and David J. Heeger. Human cortical activity correlates with stereoscopic depth perception. *Journal of Neurophysiology*, 86:2054–2068, 2001.

[BKK97]   Mitsuru Baba, Tadataka Konishi, and Nobuaki Kobayashi. A new fast rangefinding method based on a non-mechanical scanning mechanism and a high-speed image sensor. In *IEEE Instrumentation and Measurement Technology Conference*, pages 957–962, 1997.

[BMJ01]   Vladimir Brajovic, Kenichi Mori, and Nebojsa Jankovic. 100 frames/s cmos range image sensor. In *2001 IEEE International Solid State Conference*, pages 256,257 and 453, 2001.

[BMS98]   J Batlle, E Mouaddib, and J Salvi. A survey: Recent progress in coded structured light as a technique to solve the correspondence problem. *Pattern Recognition*, 31(7):963–928, 1998.

[BR86]   F. Blais and M. Rioux. Real-time numerical peak detector. In *Signal Processing*, volume 11, pages 145–155, 1986.

[CBvdHG00] Wojciech Chojnacki, Michael J. Brooks, Anton van den Hengel, and Darren Gawley. On the fitting of surfaces to data with covariances. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1294–1303, 2000.

[CBvdHG01] Wojciech Chojnacki, Michael J. Brooks, Anton van den Hengel, and Darren Gawley. A fast mle-based method for estimating

the fundamental matrix. In *Proceedings of the International Conference on Image Processing*, volume 2, pages 189–192, 2001.

[CBvdHG03] Wojciech Chojnacki, Michael J. Brooks, Anton van den Hengel, and Darren Gawley. Fns and heiv: relating two vision parameter estimation frameworks. In *Proceedings of the 12th International Conference on Image Analysis and Processing*, pages 152 – 157, 2003.

[CHJ01] Chang Wo Chu, Sungjoo Hwang, and Soon Ki Jung. Calibration-free approach to 3d reconstructuion using light stripe projections on a cube frame. In *Proceedings of the International Conference on 3D Imaging and Modelling*, pages 13–19, 2001.

[CK87] C.H. Chen and A.C. Kak. Modeling and calibration of a structured light scanner for 3-d robot vision. In *Proceedings of the IEEE conference on robotics and automation*, pages 807–815. IEEE, 1987.

[CLSB92] Guillaume Champleboux, Stéphane Lavalée, Richard Szelinski, and Lionel Brunie. From accurate imaging sensor calibration to accurate model-based 3-d object localization. In *Proceedings of the IEEE Computer Vision and Pattern Recognition conference*, pages 83–89. IEEE, 1992.

[CLSC92] G. Champleboux, S. Lavallée, P. Sautout, and P. Cinquin. Accurate calibration of cameras and range imaging sensors: the npbs method. In *Proceedings of the 1992 IEEE International Conference on Robotics and Automation*, pages 1552–1557. IEEE, 1992.

[CMRt96] E. Coste-Manière, V. Rigaud, and UNION team. Union: Underwater intelligent operation and navigation. Technical Report RR–3038, INRIA, 1996.

[Cur97] Brian Lee Curless. *New methods for surface reconstruction from range images*. PhD thesis, Stanford University, 1997.

[DC01] James Davis and Xing Chen. A laser range scanner designed for minimum calibration complexity. In *Proceedings of the International Conference on 3D Imaging and Modelling*, pages 91–98, 2001.

[dHBCG00] A. Van den Hengel, M. Brooks, W. Chojnacki, and D. Gaw-
           ley. Approximated maximum likelyhood estimation and the
           fundamental numerical scheme. Technical report, University
           of Adelaide (Australia), 2000.

[Diea]   Bart Dierickx. *CMOS image sensors. Concepts.* SPIE.

[Dieb]   Bart Dierickx. *CMOS image sensors. Limits.* SPIE.

[Die00]   Bart Dierickx. Cmos image sensors: concepts and limits. A
           short course given at Photonics West 2000, January 2000.

[DJMU85]   R. Desimone, S. J.Schein, J. Moran, and L. G. Ungerleider.
           Contour, color and shape analysis beyond the striate cortex.
           Vision Research, 25:441–452, 1985.

[Fau93]   O. Faugeras. *Three-Dimensional Computer Vision: A Geo-
           metric Viewpoint.* The MIT Press, 1993.

[Fit97]   Andrew W. J. Fitzgibbon. *Stable segmentation of 2D curves.*
           PhD thesis, University of Edinburgh, 1997.

[FK03]   Ryo Furukawa and Hiroshi Kawasaki. Interactive shape acqui-
           sition using marker attache laser projector. In *Proceedings of
           the fourth Internatonal Conference on 3D Imaging and Mod-
           elling*, 2003.

[FN91]   R.B. Fisher and D.K. Naidu. A comparison of algorithms for
           subpixel peak detection. In *Proc. 1991 British Machine Vision
           Association Conf*, pages 217–225, 1991.

[FP97]   O. Faugeras and T. Papadopoulo. Grassmann-cayley algebra
           for modeling systems of cameras and the algebraic equations
           of the manifold of trifocal tensors. Rapport de recherche 3225,
           INRIA, July 1997.

[FS02]   J Forest and J Salvi. A review of laser scanning three-
           dimensional digitisers. In *Intelligent Robots and Systems, 2002*,
           pages 73–78. IEEE/RSJ, 2002.

[FSCP04]   Josep Forest, Joaquim Salvi, Enric Cabruja, and Carles Pous.
           Laser stripe peak detector for laser scanners: a fir filter ap-
           proach. In *Proceedings of the International Conference on Pat-
           tern Recognition*, volume To Appear, 2004.

[GKC90]   Andrew Gruss, Takeo Kanade, and L. Richard Carley. *Machine Vision for Three-Dimensional Scenes*, chapter A fast lightstripe rangefinding system with smart VLSI sensor, pages 381–397. Academic Press, Inc., 1990.

[GRB+01]   Guy Godin, Marc Rioux, J. Angelo Beraldin, Marc Levoy, Luc Cournoyer, and François Blais. An assessment of laser range measurement on marble surfaces. In *5th Conference on Optical 3D Measurement Techniques*, pages 49–56, October 2001.

[GTK92]   Andrew Gruss, Shigeyuki Tada, and Takeo Kanade. A vlsi smart sensor for fast range imaging. In *Proceedings of the 1992 IEEE/RSJ International Conference on Intelligent Robots ans Systems*, pages 349–358, 1992.

[Har97]   Richard I. Hartley. Triangulation. *Computer Vision and Image Understanding*, pages 146–157, 1997.

[HBK94]   Yuko Hori, Mitsuru Baba, and Tadataka Konishi. A simultaneous high-speed measuring system of multiple points. In *IMTC'94*, pages 1365–1368, 1994.

[HJTS82]   E.L. Hall, C.A. McPherson J.B.K. Tio, and F.A. Sadhadi. Measuring curved surfaces for robot vision. *Computer Journal*, 1982.

[HK99]   E. Horn and N. Kiryati. Toward optimal structured light patterns. *Image and Vision Computing*, 17(2):87–97, February 1999.

[HS95]   Kazuyuki Hattori and Yukio Sato. Handy rangefinder for active vision. In *Proceedings of the 1995 ieee international conference on robotics and automation*, pages 1423–1428, 1995.

[HS96]   Kazuyuki Hattori and Yukio Sato. Accurate rangefinder with laser pattern shifting. In *Proceedings of the 1996 ieee international conference on robotics and automation*, pages 849–853, 1996.

[HSG00]   Khalil C. Haddad, Henry Stark, and Nikolas P. Galatsanos. Constrained fir filter design by the method of vector space projections. *IEEE Transactions on circuits and systems II: Analog and digital signal processing*, 47(8):714–725, august 2000.

[Huy97]   Du Q. Huynh. Calibration of a structured light system: a projective approach. In IEEE, editor, *Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 225–230, 1997.

[HZ03]   Richard Hartley and Andrew Zisserman. *Multiple view geometry, 2nd edition*. Cambridge University Press, 2003.

[ISM84]   S. Inokuchi, K. Sato, and F. Matsuda. Range imaging system for 3-D object recognition. In *Proceedings of the International Conference on Pattern Recognition*, pages 806–808, 1984.

[Jar83]   R. A. Jarvis. A perspective on range finding techniques for computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(2):122–139, 1983.

[Jok99]   O. Jokinen. Self-calibration of a light striping system by matching multiple 3-d profile maps. In *Proceedings of the Second International Conference on 3-D Digital Imaging and Modeling*, pages 180 – 190, 1999.

[Kan96]   Kenichi Kanatani. *Statistical optimization for geometric computation: theory and practice*. Elsevier, 1996.

[KGC91]   Takeo Kanade, A. Gruss, and L.R. Carley. A very fast vlsi rangefinder. In *Proc. IEEE International Conference on Robotics and Automation*, volume 2, pages 1322–1329, April 1991.

[KOK00]   Kenichi Kanatani, Naoya Ohta, and Yasushi Kanazawa. Optimal homography computation with a reliability measure. *IEICE Transactions on Informatics and Systems*, E38-D(7):1369–1374, 2000.

[LCM00]   L. Li, D. Cochran, and R. Martin. Target tracking with an attentive foveal sensor. In *Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 182 – 185, 2000.

[LM00]   Yoram Leedan and Peter Meer. Heteroscedastic regression in computer vision: Problems with bilinear constraint. *International Journal of Computer Vision*, 37(2):127–150, 2000.

[LRMM88]   J. Lazzaro, S. Ryckebusch, M.A. Mahowald, and C.A. Mead. Winner-take-all networks of o(n) complexity. In D.Tourestzky, editor, *Advances in Neural Information Processing Systems*, volume 1, pages 703–711, 1988.

[MT96]   Roger Mohr and Bill Triggs. Projective geometry for image analysis. Technical report, INRIA, 1996. A tutorial given at ISPRS in Vienna 1996.

[NWK88]   Kouichi Nakano, Yasuo Watanabe, and Sukeyasu Kanno. Extraction and recognition of 3-dimensional information by projecting a pair of slit-ray beams. In *Proceedings of the 9th International Conference on Pattern Recognition*, pages 736–743, 1988.

[OF86]   G. Toscani O.D. Faugeras. The calibration problem for stereo. In *Proceedings of the 1986 IEEE Computer Vision and Pattern Recognition*, pages 15–20, 1986.

[OIA03a]   Y. Oike, M. Ikeda, and K. Asada. A cmos image sensor for high-speed active range finding using column-parallel time-domain adc and position encoder. *IEEE Transactions on Electron Devices*, 50:152–158, January 2003.

[OIA03b]   Y. Oike, M. Ikeda, and K. Asada. High-speed position detector using new row-parallel architecture for fast collision prevention system. In *Proceedings of the 2003 International Symposium on Circuits and Systems*, volume 4, pages 788–791, 2003.

[PA82]   J. L. Posdamer and M. D. Altschuler. Surface measurement by space-encoded projected beam systems. *Computer Graphics and Image Processing*, 18(1):1–17, 1982.

[PBAA75]   R.J. Popplestone, C.M. Brown, and G.F. Crawford A.P. Ambler. Forming models of plane-and-cylinder faceted bodies from light stripes. In *International Joint Conferences on Artificial Intelligence*, pages 664–668, 1975.

[PDK01]   Johnny Park, Guilherme N. DeSouza, and Avinash C. Kak. Dual-beam structured-light scanning for 3d object modelling. In *Proceeidings of the 3DIM*, pages 65–72, 2001.

[RB97]   P. Rives and J. Borrelly. Underwater pipe inspection task us-
           ing visual servoing techniques. In *IEEE/RSJ Int. Conf. on
           Intelligent Robots and Systems*, volume 1, pages 63–68, 1997.

[Rei96]   Ian D. Reid. Projective calibration of a laser-stripe range
           finder. *Image and Vision Computing*, 14:659–666, 1996.

[RSB98]   P. Ridao, J. Salvi, and J Batlle. Behaviours implemented us-
           ing a vision system based on coded structured light projection.
           In *IEE Colloquium on Underwater Applications of Image Pro-
           cessing*, volume 1998/217, pages 7/1–7/6, 1998.

[SAB02]   Joaquim Salvi, Xavier Armangué, and Joan Batlle. A com-
           parative review of camera calibrating methods with accuracy
           evaluation. *Pattern Recognition*, 35(7):1617–1635, 2002.

[Sal97]   J. Salvi. *An Approach to Coded Structured Light to Obtain
           Three Dimensional Information*. PhD thesis, Universitat de
           Girona, 1997.

[Sar91]   Tapio Saramaki. Adjustable windows for the design of fir fil-
           ters - a tutorial. In *Proceedings of the 6th Mediterranean Elec-
           trotechnical Conference*, pages 28–33, May 1991.

[Sat95]   Yukio Sato. Active rangefinding and recognition with cubic-
           scope. In *Proceedings of the 2nd asian conference on computer
           vision*, pages 211–217, 1995.

[SCR00]   G. Sansoni, M. Carocci, and R. Rodella. Calibration and per-
           formance evaluation of a 3-D imaging sensor based on the pro-
           jection of structured light. *IEEE Transactions on instrumen-
           tation and measurement*, 49(3):628–636, June 2000.

[SKF82]   Yukio Sato, Hiroo Kitagawa, and Hiroichi Fujita. Shape mea-
           surement of curved objects using multiple slit-ray projections.
           *IEEE Transactions on Pattern Analysis and Machine Intel-
           lignece*, PAMI-4(6):641–646, 1982.

[SO93]   Yukio Sato and Masaki Otsuki. Three-dimensional shape re-
           construction by active rangefinder. In *Proceedings of the IEEE
           Computer Society Conference on Computer Vision and Pattern
           Recognition*, pages 142–147. IEEE, 1993.

[SPB04]   Joaquim Salvi, Jordi PagÃ¨s, and Joan Balle. Pattern codifica-
            tion strategies in structured light systems. *Pattern Recognition*,
            37:827–849, 2004.

[SS71]   Yoshiaki Shirai and Motoi Suwa. Recognition of polyhedrons
            with a range finder. In *International Joint Conferences on
            Artificial Intelligence*, pages 80–87, 1971.

[Suc91]   D. Suckley. Genetic algorithm in the design of fir filters. In
            Circuits, Devices and Systems, IEE Proceedings G, volume 138,
            pages 234–238, April 1991.

[SYI94a]   Kosuke Sato, Atsushi Yokohama, and Seiji Inokuchi. Silicon
            range finder – a realtime range finding vlsi sensor –. In *IEEE
            1994 Custom Integrated Circuits Conference*, pages 14.5.1–
            14.5.4, 1994.

[SYI94b]   Kotsuke Sato, Atsushi Yokohama, and Seiji Inokuchi. Silicon
            range finder. In *Proceedings of the IEEE 1994 custom inte-
            grated circuits conference*, pages 339–342, 1994.

[TFFN98]   E. Trucco, R.B. Fisher, A.W. Fitzgibbon, and D.K. Naidu.
            Calibration, data consistency and model acquisition with a 3-
            d laser striper. *International Journal of Computer Integrated
            Manufacturing*, 11(4):292–310, 1998.

[Tsa86]   R.Y. Tsai. *Synopsis of recent progress on camera calibration
            for 3D machine vision*, chapter The Robotic Review, pages
            147–159. The MIT Press, 1986.

[Tsa87]   R.Y. Tsai. A versatile camera calibration technique for high
            accuracy 3d machine vision metrology using off-the-shelf tv
            cameras and lenses. *IEEE International Journal on Robotics
            and Automation*, RA-3:323–344, 1987.

[TWVC99]   Masahiro Takatsuka, Geoff A.W. West, Svetha Venkatesh, and
            Terry M. Caelli. Low-cost interactive active monocular range
            finder. In *Proceeidings of the International Conference on 3D
            Imaging and Modelling*, pages 444–449, 1999.

[VdH81]   H.C. Van de Hulst. *Light Scattering by Small Particles*. Dover
            Publications, Inc., New York, 1981.

[WCH92]   J. Weng, P. Cohen, and M. Herniou. Camera calibration with
            distorsion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:965–980, 1992.

[WFAR99]   N. Werghi, R.B. Fisher, A. Ashbrook, and C. Robertson. Improving second-order surfaces estimation. In *Proceedings of the British Machine Vision Conference*, pages 275–284, 1999.

[XM02]   Ya Xue and D. Morrell. Adaptive foveal sensor for target tracking. In *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 848 – 852, 2002.

[XM03]   Ya Xue and D. Morrell. Target tracking and data fusion using multiple adaptive foveal sensors. In *Proceedings of the Sixth International Conference of Information Fusion*, volume 1, pages 326 – 333, 2003.

[YSYI94]   A. Yokoyama, K. Sato, T. Yoshigahara, and S. Inokuchi. Realtime range imaging using adjustment-free photo-vlsi. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1751–1758, 1994.

[YZWL98]   Xiaoyang Yu, Jian Zhang, Liying Wu, and Qing Lin. Calibration errors in laser scanning 3d vision measurement using the space encoding method. In *Automated optical inspection*, volume 3558, pages 298–303. SPIE, 1998.

[YZWQ96]   Xiaoyang Yu, Jian Zhang, Liying Wu, and Xifu Qiang. Laser scanning device used in space-encoding rangefinder. In *Automated optical inspection for industry*, pages 490–495. SPIE, 1996.